

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 12.9 Seconds  
(without alignments)  
16.008 million cell updates

Title: US-09-991-508B-1  
Perfect score: 26  
Sequence: 1 EGDC 4

Scoring table: ELCSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first: 45 summaries

```
Database : Issued Patents: Aa.*
1: /cgn2_6/pdata/2/iaa/5A COMB. pep.*
2: /cgn2_6/pdata/2/iaa/5B COMB. pep.*
3: /cgn2_6/pdata/2/iaa/6A COMB. pep.*
4: /cgn2_6/pdata/2/iaa/6B COMB. pep.*
5: /cgn2_6/pdata/2/iaa/PTUS COMB. pep.*
6: /cgn2_6/pdata/2/iaa/backfiles1. pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	26	100.0	4	2	US-08-335-832-41		Sequence 41, Appl
2	26	100.0	4	2	US-08-753-781-34		Sequence 34, Appl
3	26	100.0	4	3	US-09-141-127-14		Sequence 14, Appl
4	26	100.0	4	3	US-08-960-054A-1		Sequence 1, Appl
5	26	100.0	4	3	US-08-958-993A-1		Sequence 1, Appl
6	26	100.0	4	4	US-08-959-206A-4		Sequence 4, Appl
7	26	100.0	4	4	US-09-925-715-4		Sequence 4, Appl
8	26	100.0	4	6	5384309-4		Patent No. 5384309
9	26	100.0	5	1	US-08-212-186A-1C		Sequence 10, Appl
10	26	100.0	5	1	US-08-425-238-8		Sequence 8, Appl
11	26	100.0	5	1	US-08-482-880-1		Sequence 1, Appl
12	26	100.0	5	2	US-08-273-274-1		Sequence 1, Appl
13	26	100.0	5	2	US-08-475-241-1		Sequence 1, Appl
14	26	100.0	5	2	US-08-625-695A-10		Sequence 10, Appl
15	26	100.0	5	2	US-08-335-832-42		Sequence 42, Appl
16	26	100.0	5	2	US-08-753-781-35		Sequence 35, Appl
17	26	100.0	5	2	US-08-484-773-1		Sequence 1, Appl
18	26	100.0	5	2	US-08-361-864-5		Sequence 5, Appl
19	26	100.0	5	2	US-08-361-864-35		Sequence 35, Appl
20	26	100.0	5	2	US-08-286-861-37		Sequence 37, Appl
21	26	100.0	5	3	US-C9-141-127-15		Sequence 15, Appl
22	26	100.0	5	3	US-C8-924-002-10		Sequence 10, Appl
23	26	100.0	5	5	5384309-1		Patent No. 5384309
24	26	100.0	5	6	5384309-2		Patent No. 5384309
25	26	100.0	5	6	5384309-5		Patent No. 5384309
26	26	100.0	5	6	5384309-6		Patent No. 5384309
27	26	100.0	5	6	5384309-7		Patent No. 5384309

## ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

; OTHER INFORMATION: /note= "The carboxyl terminus is modified to an  
; OTHER INFORMATION: amide"  
US-08-335-632-41

Query Match 100.0%; Score 26; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 2

US-08-753-781-34  
; Sequence 34, Application US/03753781C  
; Patent No. 5951981  
; GENERAL INFORMATION:  
; APPLICANT: Markland Jr., Francis S.  
; APPLICANT: Bush, Larry R.  
; APPLICANT: Swenson, Stephen  
; APPLICANT: Flores Sanchez, Eladio  
; TITLE OF INVENTION: THROMBOLYTIC AGENTS WITH ANTITHROMBOTIC ACTIVITY  
; FILE REFERENCE: DITI 124  
; CURRENT APPLICATION NUMBER: US/08/753,781C  
; CURRENT FILING DATE: 1996-12-02  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: ACETYLTATION  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (4)  
; OTHER INFORMATION: AMIDATION  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic  
; OTHER INFORMATION: peptide  
US-08-753-781-34

Query Match 100.0%; Score 26; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 3

US-09-141-127-14  
; Sequence 14, Application US/09141127A  
; Patent No. 6083481  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T.  
; APPLICANT: Lister-James, John  
; TITLE OF INVENTION: THROMBUS IMAGING AGENTS  
; FILE REFERENCE: DITI 113.1USC1  
; CURRENT APPLICATION NUMBER: US/09/141,127A  
; CURRENT FILING DATE: 1998-08-27  
; EARLIER APPLICATION NUMBER: 08/335,832  
; EARLIER FILING DATE: 1995-01-05  
; EARLIER APPLICATION NUMBER: PCT/US93/04794  
; EARLIER FILING DATE: 1993-05-21  
; EARLIER APPLICATION NUMBER: 07/886,752  
; EARLIER FILING DATE: 1992-05-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: ACETYLTATION  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (4)  
; OTHER INFORMATION: AMIDATION  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic  
; OTHER INFORMATION: peptide  
US-09-141-127-14

Query Match 100.0%; Score 26; DB 3; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 4

US-08-960-054A-1  
; Sequence 1, Application US/08960054A  
; Patent No. 6261537  
; GENERAL INFORMATION:  
; APPLICANT: Nycomed Imaging AS  
; TITLE OF INVENTION: Improvements in or relating to  
; TITLE OF INVENTION: diagnostic/therapeutic  
; FILE REFERENCE: REF/Klaveness/054  
; CURRENT APPLICATION NUMBER: US/08/960,054A  
; CURRENT FILING DATE: 1997-10-29  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial  
; OTHER INFORMATION: Sequence:RGDC-Mal-PEG3400-DSPE  
US-08-960-054A-1

Query Match 100.0%; Score 26; DB 3; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 5

US-08-958-993A-1  
; Sequence 1, Application US/08958993A  
; Patent No. 6264917  
; GENERAL INFORMATION:  
; APPLICANT: Nycomed Imaging AS  
; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic  
; TITLE OF INVENTION: agents  
; FILE REFERENCE: REF/Klaveness/993  
; CURRENT APPLICATION NUMBER: US/08/958,993A  
; CURRENT FILING DATE: 1997-10-28  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
US-08-958-993A-1
; OTHER INFORMATION: Sequence:RGDC-Mal-PEG3400-DSPE

Query Match      100.0%; Score 26; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      1 RGDC 4

RESULT 6
US-08-959-206A-4
; Sequence 4, Application US/08959206A
; Patent No. 6331289
; GENERAL INFORMATION:
; APPLICANT: Nycomed Imaging AS
; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic
; FILE REFERENCE: REF/Klaveness/206
; CURRENT APPLICATION NUMBER: US/08/959.206A
; CURRENT FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic RGDC
US-08-959-206A-4

Query Match      100.0%; Score 26; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      1 RGDC 4

RESULT 7
US-09-925-715-4
; Sequence 4, Application US/09925715
; Patent No. 6683047
; GENERAL INFORMATION:
; APPLICANT: Nycomed Imaging AS
; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic
; FILE REFERENCE: REF/Klaveness/206
; CURRENT APPLICATION NUMBER: US/09/925.715
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic RGDC
US-09-925-715-4

Query Match      100.0%; Score 26; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      1 RGDC 4

```

```

QY      1 RGDC 4
Db      1 RGDC 4

RESULT 8
5384309-4
; Patent No. 5384309
; APPLICANT: BARKER, PETER L.;BURNIER, JOHN P.;THORSETT, EUGENE D.
; TITLE OF INVENTION: CYCLIZED PEPTIDES AND THEIR USE AS PLATELET
; AGGREGATION INHIBITORS
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/380,957
; FILING DATE: 17-JUL-1989
; SEQ ID NO:4
; LENGTH: 4
5384309-4

Query Match      100.0%; Score 26; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      1 RGDC 4

RESULT 9
US-08-212-186A-10
; Sequence 10, Application US/08212186A
; Patent No. 5536814
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki I.
; APPLICANT: Koivunen, Erkki
; TITLE OF INVENTION: NOVEL INTEGRIN-BINDING PEPTIDES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,186A
; FILING DATE: 11-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: both
US-08-212-186A-10

Query Match      100.0%; Score 26; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      1 RGDC 4

```

Db 2 RGDC 5

RESULT 10

US-08-425-238-8

Sequence 8, Application US/08425238

Patent No. 5627263

GENERAL INFORMATION:

APPLICANT: Ruoslahti, Erkki

APPLICANT: Koivinen, Erkki

TITLE OF INVENTION: No. 5627263el Integrin-Binding Peptides

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,238

FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/158,001

FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LA 9775

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-425-238-8

Query Match 100.0%; Score 26; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4

Db 2 RGDC 5

RESULT 11

US-08-482-880-1

Sequence 1, Application US/08482880

Patent No. 5736122

GENERAL INFORMATION:

APPLICANT: Dean, Richard T

APPLICANT: Lister-James, John

TITLE OF INVENTION: Technetium-99m Labeled Peptides for

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 10 South Wacker Drive Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,880

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 5736122han, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 92,216-L

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Cross-links

LOCATION: 1..5

OTHER INFORMATION: /label= thioether

OTHER INFORMATION: /note= "The sidechain sulfur of the C-terminal

OTHER INFORMATION: cysteine residue is covalently linked as an ether

OTHER INFORMATION: to the group [CH2CO], which group forms an amide

US-08-482-880-1

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,880

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 5736122han, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 92,216-L

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Cross-links

LOCATION: 1..5

OTHER INFORMATION: /label= thioether

OTHER INFORMATION: /note= "The sidechain sulfur of the C-terminal

OTHER INFORMATION: cysteine residue is covalently linked as an ether

OTHER INFORMATION: to the group [CH2CO], which group forms an amide

US-08-482-880-1

Query Match 100.0%; Score 26; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4

Db 2 RGDC 5

RESULT 12

US-08-273-274-1

Sequence 1, Application US/08273274

Patent No. 5849260

GENERAL INFORMATION:

APPLICANT: Dean, Richard T

APPLICANT: Lister-James, John

TITLE OF INVENTION: Technetium-99m Labeled Peptides for

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/273,274

FILING DATE:

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/886,752

FILING DATE: 21-MAY-1992

ATTORNEY/AGENT INFORMATION:

NAME: No. 5849260nac, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 92,216



```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Cross-links
; LOCATION: 1..5
; OTHER INFORMATION: /label= thioether
; OTHER INFORMATION: /note= "The sidechain sulfur of the C-terminal
; OTHER INFORMATION: cysteine residue is covalently linked as an ether
; OTHER INFORMATION: to the group [CH2CO], which group forms an amide
US-08-273-274-1

Query Match 100.0%; Score 26; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 2 RGDC 5

RESULT 13
US-08-475-041-1
; Sequence 1, Application US/08475041
; Patent No. 5879658
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Technetium-99m Labeled Peptides for
; TITLE OF INVENTION: Thrombus Imaging
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,041
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5879658-Jan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,216-M
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Cross-links
; LOCATION: 1..5
; OTHER INFORMATION: /label= thioether
; OTHER INFORMATION: /note= "The sidechain sulfur of the C-terminal
```

```
; OTHER INFORMATION: cysteine residue is covalently linked as an ether
; OTHER INFORMATION: to the group [CH2CO], which group forms an amide
US-08-475-041-1

Query Match 100.0%; Score 26; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 2 RGDC 5

RESULT 14
US-08-625-695A-10
; Sequence 10, Application US/08625695A
; Patent No. 5912234
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki I.
; APPLICANT: Koivunen, Erkki I.
; TITLE OF INVENTION: NOVEL INTEGRIN-BINDING PEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,695A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,186
; FILING DATE: 11-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-625-695A-10

Query Match 100.0%; Score 26; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 2 RGDC 5

RESULT 15
US-08-335-832-42
; Sequence 42, Application US/08335832
; Patent No. 5925331
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Technetium-99m Labeled Peptides for
; TITLE OF INVENTION: Thrombus Imaging
```

```

: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Bandex & Allegretti, Ltd.
: STREET: 10 South Wacker Drive Suite 3000
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/335,832
: FILING DATE: 05-JAN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Mr. 592531nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 92,216-I
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 42:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-335-832-42

```

```

Query Match      100.3%; Score 26; DB 2; Length 5;
Best Local Similarity 100.3%; Pred.No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qv      1 RGDC 4
Db      2 RGDC 5

```

```

Search completed: April 16, 2004, 07:31:37
Job time : 13.9 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:55:57 ; Search time 43.7 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588B-1

Perfect score: 26

Sequence: 1 RGDC 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	4	2 AAR11744	Aar11744 Cyclic pl
2	26	100.0	4	2 AAR242-4	Aa24214 Fragment
3	26	100.0	4	2 AAR69324	Aa69324 Gp Iib/II
4	26	100.0	4	2 AAW25175	Aa25175 RGD-pepti
5	26	100.0	4	2 AAW64951	Aa64951 Targettin
6	26	100.0	4	3 AAY52481	Aay52481 Fibrinect
7	26	100.0	4	3 AAY80468	Aay80468 Cell adhe
8	26	100.0	4	3 AAY95464	Aay95464 GPIIb/III
9	26	100.0	4	4 AAB91949	Aab91949 Fibrinect
10	26	100.0	4	4 AAB86860	Aab86860 Transport
11	26	100.0	4	6 AAP60201	Aap60201 Cell adhe
12	26	100.0	4	7 ADC83765	Adc83765 Carboxyal
13	26	100.0	4	7 ADC77589	Adc77589 Cell adhe
14	26	100.0	5	2 AAR11587	Aar11587 Fibrinoge
15	26	100.0	5	2 AAR10415	Aar10415 Fibrinoge
16	26	100.0	5	2 AAR10413	Aar10413 Fibrinoge
17	26	100.0	5	2 AAR10414	Aar10414 Fibrinoge
18	26	100.0	5	2 AAR10418	Aar10418 Fibrinoge
19	26	100.0	5	2 AAR10770	Aar10770 Platelet
20	26	100.0	5	2 AAR10773	Aar10773 Platelet
21	26	100.0	5	2 AAR10768	Aar10768 Platelet
22	26	100.0	5	2 AAR10769	Aar10769 Platelet
23	26	100.0	5	2 AAR10771	Aar10771 Platelet
24	26	100.0	5	2 AAR10772	Aar10772 Platelet
25	26	100.0	5	2 AAR28190	Aar28190 Peptide c

26	26	100.0	5	2 AAR28204	Aar28204 Peptide c
27	26	100.0	5	2 AAR27031	Aar27031 Peptide l
28	26	100.0	5	2 AAR29052	Aar29052 Peptide l
29	26	100.0	5	2 AAR69325	Aar69325 Gp Iib/II
30	26	100.0	5	2 AAR79093	Aar79093 Alpha5/be
31	26	100.0	5	2 AAW03492	Aaw03492 Alpha(5) -
32	26	100.0	5	2 AAW79644	Aaw79644 Cyclo(S,S
33	26	100.0	5	2 AAW48499	Aaw48499 Integrin
34	26	100.0	5	2 AAW64952	Aaw64952 Targettin
35	26	100.0	5	2 AAW50594	Aaw50594 GPIIb/III
36	26	100.0	5	2 AAY21570	Aay21570 Integrin-
37	26	100.0	5	2 ADE25437	Ade25437 GPIIb/III
38	26	100.0	5	2 ADE25491	Ade25491 Tc-99m la
39	26	100.0	5	3 AAY54931	Aay54931 Peptide l
40	26	100.0	5	3 AAY54976	Aay54976 Peptide l
41	26	100.0	5	3 AAY95465	Aay95465 GPIIb/III
42	26	100.0	5	4 AAB82080	Aab82080 Peptidic
43	26	100.0	5	4 AAB83205	Aab83205 Integrin
44	26	100.0	5	5 AAE17982	Aae17982 Human lig
45	26	100.0	5	5 ABG30410	Abg30410 N,N'-dist

ALIGNMENTS

RESULT 1  
AAR11744  
ID AAR11744 standard, peptide; 4 AA.  
XX AC AAR11744;  
XX AC  
XX 25-MAR-2003 (revised)  
DT 05-JUL-1991 (first entry)  
XX  
DE Cyclic platelet aggregation inhibitor - example #1.  
XX MYocardial infarction; deep vein thrombosis; pulmonary embolism; stroke.  
OS Synthetic.  
XX  
XX EP425212-A.  
XX  
XX C2-MAY-1991.  
XX  
XX 22-OCT-1990; 90BP-00311537.  
XX  
XX 23-OCT-1989; 89US-00425906.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM.  
XX (ALIF/) ALI F E.  
XX Ali FEF, Samanen JM;  
XX WPI, 1991-126685/18.  
XX  
XX New cyclic peptide(s) with Gly-Asp sequence - useful as platelet aggregation inhibitors to treat myocardial infarction, deep vein thrombosis, pulmonary embolism, stroke etc.  
XX Claim 9; Page 44; 55pp; English.

The sequence is the peptide part of the cpd.: cyclo (S,S)-Mba-Arg-Gly-Asp -Cys-NH2 Mba-2-mercapto-bezoic acid. See also AAR11743-RI1748 for examples. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

Sequence 4 AA;

Query Match 100.0%; Score 26; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4

DE 1 RGDC 4

RESULT 2  
AAR24214  
ID AAR24214 standard; protein; 4 AA.  
XX  
AC AAR24214;  
XX  
DT 25-MAR-2003 (revised)  
DT 09-JAN-2003 (revised)  
DT 18-NOV-1992 (first entry)  
XX  
DE Fragment of tenascin-related peptide.  
XX  
KW Tenascin; related peptide; cell attachment; antibody; angiogenesis;  
KW tumour metastasis; solid matrix; prosthetic device; vascular graft;  
KW percutaneous device.  
XX  
OS Synthetic.  
XX  
PN WO9207872-A1.  
XX  
PD 14-MAY-1992.  
XX  
PF 29-OCT-1991; 91WO-US08018.  
XX  
PR 29-OCT-1990; 90US-00605920.  
PR 30-OCT-1990; 90US-00605667.  
XX  
PA (CALB-) CALIFORNIA INST BIOLOGICAL RES.  
XX  
PI Bourdon MA;  
XX  
DR WPI; 1992-183625/22.  
XX  
PT New tenascin-related peptides - modulate cell attachment to tenascin,  
PT useful in inhibition of tumour metastasis and angiogenesis.  
XX  
PS Disclosure; Page 8; 60pp; English.  
XX  
CC The peptide may form an N- or C-terminal fragment of the generic peptide  
CC of AAR24192, which is a tenascin-related peptide. This peptide mimics the  
CC ability of tenascin to promote cell attachment. The peptide and  
CC antibodies raised to it can be used to modulate cell attachment to  
CC tenascin, esp. to inhibit tumour metastasis and angiogenesis. The peptide  
CC is pref. attached to a solid matrix, eg collagen, nitrocellulose,  
CC polyester, glass, synthetic resin, long-chain polysaccharide or synthetic  
CC resin fibre. It is esp. operatively linked to a solid matrix forming a  
CC prosthetic device, percutaneous device, vascular graft, etc. For topical  
CC admin. it is formulated into a lotion, salve, gel, colloid, powder etc.  
CC (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 26; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+36;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDC 4  
Db 1 RGDC 4  
RESULT 3  
AAR69324  
ID AAR69324 standard; peptide; 4 AA.  
XX  
AC AAR69324;  
XX  
DT 25-MAR-2003 (revised)  
XX

DT 25-JUN-1995 (first entry)  
XX  
DE Gp IIB/IIIA receptor ligand used in scintigraphic imaging of thrombi.  
XX  
KW Scintigraphy; thrombus; thrombi; imaging; specific binding;  
KW technetium-99m; radiolabelled; Gp IIB/IIIA receptor ligand.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "acetyl-Arg"  
FT Modified-site 4  
FT /note= "Cys-amide"  
XX  
PN WO9323085-A1.  
XX  
PD 25-NOV-1993.  
XX  
PF 21-MAY-1993; 93WO-US004794.  
XX  
PR 21-MAY-1992; 92US-00806752.  
XX  
PA (DIAT-) DIATECH INC.  
XX  
PI Dean RT, Lister-James J;  
XX  
DR WPI; 1993-386225/48.  
XX  
PT Reagent for scintigraphic imaging of thrombi with 99m technetium -  
PT comprises synthetic peptide which binds to thrombus covalently coupled to  
PT metal binding gp., rapidly cleared from blood and tissue.  
XX  
PS Claim 41; Page 46; 61pp; English.  
XX  
CC The invention relates to reagents for scintigraphic imaging of a thrombus  
CC in-vivo, comprising (A) a specific binding compound capable of binding to  
CC at least one component of a thrombus, covalently linked to (B) a  
CC technetium-99m-binding moiety. Specific peptides constituting the  
CC reagents are claimed as new. The present peptide is one such peptide.  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 26; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDC 4  
Db 1 RGDC 4  
RESULT 4  
AAW25175  
ID AAW25175 standard; peptide; 4 AA.  
XX  
AC AAW25175;  
XX  
DT 05-JAN-1998 (first entry)  
XX  
DE RGD-peptide capable of binding cell adhesion molecules.  
XX  
KW RGD; arginine; glycine; aspartic acid; cell adhesion molecule; binding;  
KW bladder irrigation; tumour removal; endoscopic operation;  
KW transurethral resection; cancer; neoplasia.  
XX  
OS Synthetic.  
XX  
PN DE19529909-A1.  
XX  
PD 20-FEB-1997.  
XX

PF 15-AUG-1995; 95DB-01028909.  
 XX  
 PR 15-AUG-1995; 95DB-01028909.  
 XX  
 XX (PREP ) PRESENIUS AG.  
 XX  
 XX Boehle A;  
 PI  
 XX WPI; 1997-133793/13.  
 DR  
 XX Endoscopic irrigation sclns. - contg. peptide(s) that bind to cell  
 PT adhesion molecules.  
 PT  
 XX Claim 5; Page 8; 8pp; German.  
 PS  
 XX AAW25173-W25186 are peptides containing an RGD sequence or equivalent.  
 CC The peptides are capable of binding to cell adhesion molecules and are  
 CC used in aqueous irrigation solutions for use during and after endoscopic  
 CC operations. Preferred irrigation solutions are electrolyte-free and  
 CC contain 1 microlit/ml to 100 mg/ml of one or more oligopeptides containing  
 CC the amino acid sequences: RGD, LDV, IDA, DGEA, GRRP, VTL, YIGSR, KOAGDV  
 CC and/or REDV (Given in one letter amino acid code). The solutions are  
 CC especially used for irrigating the bladder during and after tumour  
 CC removal by transurethral resection. The peptides protect against  
 CC recurrence of tumours  
 CC  
 XX Sequence 4 AA;  
 SQ  
 Query Match 100.0%; Score 26; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 1 RGDC 4  
 RESULT 5  
 AAW64951  
 ID AAW64951 standard; peptide; 4 AA.  
 AC  
 AC AAW64951;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-SEP-1998 (first entry)  
 DE  
 DE Targeting peptide #44 useful as component of thrombolytic agent.  
 XX  
 KW Thrombolytic agent; thrombolytic proteinase; blood clot; fibrin;  
 KW thrombus; antithrombotic activity.  
 XX  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 4 /note= "C-terminal amide"  
 FT  
 FT W09824917-A1.  
 XX  
 XX PD 11-JUN-1998  
 XX  
 PF 02-DEC-1997; 97WO-US021918.  
 XX  
 XX 02-DEC-1996; 96US-00753781.  
 XX (DIAT-) DIATIDE INC.  
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX Mark-and FS, Bush LR, Swenson S, Flores Sanchez E;  
 XX WPI; 1998-333336/29.  
 XX  
 XX New thrombolytic agents - comprise thrombolytic proteinase covalently  
 linked to targeting compound for binding to component of thrombus.  
 XX  
 XX Claim 10; Page 65; 79pp; English.  
 PS  
 XX The invention relates to new thrombolytic agents which comprise a  
 CC thrombolytic proteinase covalently linked to a targeting compound  
 CC capable of specifically binding to a component of a thrombus. The  
 CC thrombolytic agents can be used for eliminating thrombi in vivo in, e.g.  
 CC myocardial infarction, cerebral ischaemia, deep vein thrombosis or  
 CC pulmonary embolism. A labelled form of the thrombolytic agent can also be  
 CC used to image thrombi for diagnostic purposes. The thrombolytic agents  
 CC are specifically targeted to thrombus sites in vivo and have minimal  
 CC haemorrhagic side effects and side effects related to non-specific  
 CC proteolysis. The present sequence represents a specifically claimed  
 CC targeting peptide. (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 26; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 1 RGDC 4  
 RESULT 6  
 AAY52481  
 ID AAY52481 standard; peptide; 4 AA.  
 AC  
 AC AAY52481;  
 XX  
 DT 22-FEB-2000 (first entry)  
 DE  
 DE Fibrinectin-derived tetrapeptide #2 with cell attachment activity.  
 XX  
 KW Cell attachment; adhesion; fibronectin; tetrapeptide; collagen;  
 KW fibrinogen; cell surface protein; recognition; substrate; universal;  
 KW phagocytosis; infection; bacterial; viral; cell culture; prosthesis;  
 KW injury; treatment; reconstruction; inhibition.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "One or more amino acids may be optionally  
 FT conjugated to the N-terminus, or may be absent"  
 FT Modified-site 4 /note= "One or more amino acids may be optionally  
 FT conjugated to the C-terminus, or may be absent"  
 FT  
 FT US5394501-A.  
 XX  
 XX PD 30-NOV-1999.  
 XX  
 XX 10-APR-1998; 98US-00058540.  
 XX  
 XX 04-AUG-1982; 82US-00405239.  
 PR 08-OCT-1982; 82US-00433457.  
 PR 28-JUL-1983; 83US-00518036.  
 PR 22-NOV-1983; 83US-00554821.  
 PR 17-JUN-1985; 85US-00744981.  
 PR 09-SEP-1988; 88US-00242713.  
 PR 25-FEB-1991; 91US-00660526.  
 PR 19-JUN-1992; 92US-00902472.  
 PR 02-JUN-1995; 95US-00458833.  
 XX (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX

PI Pierschbacher M, Ruoslahti E;  
 XX WPI; 2000-052346/04.  
 XX  
 XX  
 PT Tetrapeptides derived from fibronectin useful for attaching biological  
 FT matter, especially cells, to substrates.  
 XX  
 XX  
 ES Claim 1; Col 2; 8pp; English.  
 XX  
 XX This sequence represents a fibronectin-derived tetrapeptide (#2) which  
 CC has cell attachment activity. Fibronectin promotes the attachment of  
 CC suspended cells to collagen, and also promotes the attachment of  
 CC suspended cells directly to tissue culture substrate, independent of its  
 CC binding to collagen. This sequence, or a chemically similar, biologically  
 CC equivalent sequence is shared by some other proteins that interact with  
 CC cells, including collagen, fibrinogen, and surface proteins of E. coli  
 CC and Sindbis virus. This may indicate that there is a general recognition  
 CC system that cells use to adhere to many substrates, and that this  
 CC mechanism may also be involved in a cell's phagocytic activity. In  
 CC addition, it is possible that bacteria and certain viruses utilise this  
 CC universal cellular adhesion mechanism to gain entry into a cell. The  
 CC tetrapeptide, or peptides containing the tetrapeptide, may be used to  
 CC prepare substrates to which cells will attach. Such substrates are useful  
 CC in cell culture dishes and for use in medical prosthetic devices for  
 CC implantation into the human body where enhanced cell attachment is  
 CC desirable. This is particularly useful for therapeutic reconstruction and  
 CC treatment of injuries. Conversely, the tetrapeptide may be used in a  
 CC solubilised or suspended form to inhibit undesirable cell attachment to a  
 CC substrate or to each other and to enhance phagocytic activity of cells  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 26; DB 3; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 Db 1 RGDC 4  
 RESULT 7  
 AAY80468  
 ID AAY80468 standard; peptide; 4 AA.  
 AC AAY80468;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Cell adhesion peptide #3.  
 XX  
 KW Bone regenerative; osteopathic; osseous tissue; reconstruction;  
 KW scaffold matrix; bone formation promoter; bone resorption inhibitor;  
 KW cell adhesion; osteoblast; osteoclast; bone defect; fracture.  
 XX  
 OS Synthetic.  
 XX  
 PN W0200004941-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 22-JUL-1998; 93WO-US016800.  
 XX  
 PR 24-JUL-1998; 98US-00122348.  
 XX  
 XX (PHAR-) PHARMACAL BIOTECHNOLOGIES INC.  
 PA  
 XX Budny JA;  
 XX  
 XX WPI; 2000-195084/17.  
 XX  
 XX System for reconstructing osseous tissue, useful e.g. for treating  
 XX fractures, comprises scaffold containing promoter of bone formation and

PT inhibitor of bone resorption.  
 XX  
 XX Claim 14; Page 31; 44pp; English.  
 XX  
 CC The invention relates to a novel system for reconstruction of osseous  
 CC tissue comprising a scaffold carrying a compound (I) that promotes bone  
 CC formation and a component that decreases bone resorption (II). (I) induces  
 CC migration and adhesion of osteoblasts and osteoclasts and (II) inhibits  
 CC proteolysis (specifically by plasmin) of extracellular matrix.  
 CC (I) is preferably selected from: selectin or selectin binding fragments,  
 CC proteins and peptides that facilitate cell adhesion, plasminogen  
 CC activator inhibitors, protease inhibitors and metalloprotease inhibitors.  
 CC The peptides AAY80466-Y80492 are claimed examples of cell adhesion  
 CC peptides used in the system of the invention. The system is used to  
 CC replace, remodel or correct bone defects, e.g. fractures, fissures or  
 CC bone mass loss. Incorporation of (I) into the scaffold results in rapid  
 CC seeding by osteoblasts and the development of an organic matrix, i.e. the  
 CC preformed scaffold replaces the rate-determining step of extracellular  
 CC matrix formation. The scaffold can be designed to have a predetermined  
 CC resorption/degradation rate, and may include regulatory compounds for  
 CC specific cell types  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 26; DB 3; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 Db 1 RGDC 4  
 RESULT 8  
 AAY95464  
 ID AAY95464 standard; peptide; 4 AA.  
 XX  
 AC AAY95464;  
 XX  
 DT 10-OCT-2000 (first entry)  
 XX  
 DE GPIIb/IIIa receptor ligand.  
 XX  
 KW Thrombus; imaging; radioimaging; thrombosis; thromboembolism; embolism;  
 KW diagnosis; technetium-99m; GPIIb/IIIa receptor ligand.  
 XX  
 CS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 4 /note= "C-terminal amide"  
 FT  
 XX US6083481-A.  
 XX  
 XX 04-JUL-2000.  
 XX  
 PF 27-AUG-1998; 98US-00141127.  
 XX  
 PR 21-MAY-1992; 92US-00886752.  
 PR 21-MAY-1993; 93WO-US004794.  
 PR 05-JAN-1995; 95US-00335832.  
 XX  
 PA (DIAT-) DIATIDE INC.  
 XX  
 XX Lister-James J, Dean RT;  
 XX WPI; 2000-498061/44.  
 XX  
 XX Composition comprises technetium-99m and polyamide reagent which binds to  
 XX thrombi, useful as scintigraphic imaging agent for imaging sites of  
 XX thrombus formation in vivo.

XX Example 2; Col 15-16; 27pp; English.

PS The present sequence is that of a peptide that acts as a ligand for the

CC GPIIb/IIIa receptor. The invention relates to radiolabeled reagents that

CC are scintigraphic imaging agents for imaging sites of thrombus formation

CC in vivo. The reagents each comprise a specific binding compound, such as

CC the present peptide, that is capable of binding to at least one component

CC of a thrombus, and which is covalently linked to a radiolabel-binding

CC moiety. A method for using such a reagent labeled with technetium-99m to

CC image a thrombus site in a mammalian body is claimed

XX Sequence 4 AA;

SQ

Query Match 100.0%; Score 26; DB 3; Length 4;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4

Db 1 RGDC 4

RESULT 9

AAB91949

ID AAB91949 standard; peptide; 4 AA.

AC AAB91949;

XX 22-JUN-2001 (first entry)

DT

DE Fibrinectin fragment and fibrin related peptide SEQ ID NO:1125.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;

KW blood component; modification; succinimidy; maleimido group; amino;

KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

PN 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US013576.

PF 17-MAY-1999; 99US-0134406P.

PR 10-SEP-1999; 99US-0153406P.

PR 15-OCT-1999; 99US-0159783P.

XX (CONJ-) CONJUCHEM INC.

PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;

PI WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents

PT peptidase degradation, useful for increasing length of in vivo activity.

XX Disclosure; Page 563; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)

CC comprising a therapeutically active amino acid region (III) and a

CC reactive group (II) (e.g. succinimidy and maleimido groups) attached to

CC a less therapeutically active amino acid region (IV), which covalently

CC bonds with amino/hydroxy/thiol groups on blood components to form a

CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.

CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth

CC factors and neurotransmitters; to protect them from peptidase activity in

CC vivo for the treatment of various disorders. Endogenous therapeutic

CC peptides are not suitable as drug candidates as they require frequent

CC administration due to rapid degradation by peptidases in the body.

CC Modifying and attaching therapeutic peptides to albumin prevents or

CC reduces the action of peptidases to increase length of activity (half

CC life) and specificity as bonding to large molecules decreases

CC intracellular uptake and interference with physiological processes.

CC AAB90829 to AAB92441 represent peptides which can be used in the

CC exemplification of the present invention

XX Sequence 4 AA;

SQ

Query Match 100.0%; Score 26; DB 4; Length 4;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4

Db 1 RGDC 4

RESULT 10

AAB86860

ID AAB86860 standard; peptide; 4 AA.

XX AAB86860;

AC AAB86860;

XX 28-NOV-2001 (first entry)

DT

DE Transport molecule/ligand binding-associated peptide #6.

XX Transport molecule; ligand; cancer treatment; autoimmune disease;

KW inflammation; infection.

XX Synthetic.

XX WO200168142-A1.

PN 20-SEP-2001.

PR 13-MAR-2001; 2001WO-EP002833.

PR 13-MAR-2000; 2000DE-01012120.

XX (KTEC-) KTB TUMORFORSCHUNGS GMBH.

PA Kratz F;

PI WPI; 2001-589998/66.

XX New ligand, comprising therapeutic or diagnostic agent bonded non-

PT covalently with substance having high affinity to transport molecule.

XX Disclosure; Page 39; 74pp; German.

XX This invention describes novel ligands which bind to transport molecules,

CC comprising a therapeutic and/or diagnostic agent (A) non-covalently

CC bonded via a linkage cleavable in vivo depending on pH and/or

CC enzymatically with a substance (B) having an association constant KA to a

CC transport molecule of above  $10^3 \text{ M}^{-1}$ , is new. The medicaments are

CC especially useful for the treatment of cancers, autoimmune diseases,

CC acute and chronic inflammation and infections caused by viruses or

CC microorganisms. The diagnostic kits are useful for the detection of these

CC illnesses and for the detection of the transport molecule and/or its

CC distribution in vivo. The ligands have excellent solubility in the medium

CC at the site of action and are easy and inexpensive to convert into

CC adducts, as the interaction with the transport material is physical.

CC AAB86843-AAB86920 represent peptides used to illustrate the method of the

CC invention

XX Sequence 4 AA;

SQ

Query Match 100.0%; Score 26; DB 4; Length 4;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4

```

Db          1 RGDC 4
|||||
RESULT 11
ID ABP60201 standard; peptide; 4 AA.
XX
XX AC ABP60201;
XX
XX DT 14-FEB-2003 (first entry)
XX
XX DE Cell adhesive peptide.
XX
XX KW Substrate specificity; enzyme activity; microarray; diagnosis; cancer;
XX differential analysis; high throughput analysis; cell-adhesive peptide.
XX
XX OS Synthetic.
XX
XX PN WO200283933-A2.
XX
XX PD 24-OCT-2002.
XX
XX PF 17-APR-2002; 2002WG-EP004265.
XX
XX PR 17-APR-2001; 2001DE-C1016774.
XX
XX PA (JERI-) JERINI AG.
XX
XX PI Schneider-Mergener J, Schutkowski M, Reimer U, Dong L, Pause S;
XX Scharn D, Osterkamp F, Hummel G, Jobron L;
XX
XX DR WPI; 2003-075559/07.
XX
XX PT Determining substrate specificity of enzymes, useful e.g. in screening
XX for modulators, by detecting molecular weight changes in ordered array of
XX amino acid sequences.
XX
XX PS Disclosure; Page 2; 79pp; German.
XX
XX CC The invention relates to determining the substrate specificity of an
XX enzymatic activity (EA) by contact and/or incubation of EA with an array
XX of many amino acid sequences (I) immobilised, in a directed manner, on
XX the flat surface of a carrier then detecting reaction of (I) with EA,
XX where the reaction is monitored from a change in the molecular weight of
XX (I). The method is used to determine the pattern of EA in a sample, e.g.
XX for differential analysis; for identification of a strain or species and
XX for diagnosis of disease, e.g. cancer. It can also be used for selection
XX of active agents, i.e. compounds that modulate EA. The method is suitable
XX for high throughput testing and the use of a non-porous carrier surface
XX makes it possible to use extremely small quantities of EA or sample. The
XX signal-to-noise ratio is much lower (typically by a factor of 3000) than
XX that in conventional peptide/protein arrays. The present sequence is that
XX of a cell-adhesive peptide disclosed in the invention.
XX
XX SQ Sequence 4 AA;
Query Match 100.0%; Score 26; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 RGDC 4
DB 1 RGDC 4
|||||
RESULT 12
ADC83765
ID ADC83765 standard; peptide; 4 AA.
XX
XX AC ADC83765;
XX
XX DT 01-JAN-2004 (first entry);

```

```

XX Carboxyalkyl-tetraazacyclododecane conjugate peptide #1.
DE
XX
XX KW 1,4,7-tris-(Carboxyalkyl)-1,4,7,10-tetraazacyclododecane; NMR diagnosis;
XX radiodiagnosis; radiotherapy; tumour; cardiac infarction.
XX
XX OS Synthetic.
XX
XX PN WO2003013617-A2.
XX
XX PD 20-FEB-2003.
XX
XX PF 18-JUL-2002; 2002WO-EP008000.
XX
XX PR 20-JUL-2001; 2001DE-01035355.
XX
XX PA (SCED ) SCHERING AG.
XX
XX PI Platzek J, Schmitt-Willich H, Michl G, Freczel T, Suelzle D;
XX Bader H, Raduechel B, Weinmann H, Schirmer H;
XX
XX DR WPI; 2003-354398/33.
XX
XX CC New 1,4,7-tris-(carboxyalkyl)-1,4,7,10-tetraazacyclododecane derivative
XX biomolecule conjugates, useful in radiodiagnosis or radiotherapy or
XX especially as NMR contrast agents.
XX
XX PS Example 30-90; Page 81; 93pp; German.
XX
XX CC This invention describes novel 1,4,7-tris-(Carboxyalkyl)-1,4,7,10-
XX tetraazacyclododecane derivatives (including metal complexes), conjugated
XX with a biomolecule via a spacer. The derivatives are used as agents for
XX NMR diagnosis, radiodiagnosis or radiotherapy. Typically the derivatives
XX are used as NMR contrast agents for detecting tumours or cardiac
XX infarction or in the radiotherapy of malignant tumours. NMR contrast
XX agents with high and finely tunable relaxivity may be obtained by using
XX the special ligands in the derivatives and the biomolecule component has
XX a specific targeting effect (allowing more accurate imaging and use of
XX reduced dosages). The derivatives are stable in vitro and in vivo, have
XX high water-solubility and low osmolality (allowing use as highly
XX concentrated solutions) and are well tolerated and completely excreted
XX from the body. This sequence represents a peptide used to conjugate with
XX the derivatives of the invention.
XX
XX SQ Sequence 4 AA;
Query Match 100.0%; Score 26; DB 7; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDC 4
DB 1 RGDC 4
|||||
RESULT 13
ADC77589
ID ADC77589 standard; peptide; 4 AA.
XX
XX AC ADC77589;
XX
XX DT 01-JAN-2004 (first entry)
XX
XX DE Cell adhesive peptide.
XX
XX KW immobilization; nitrene; chemoselective; biochip.
XX
XX OS Unidentified.
XX
XX PN WO2003014139-A2.
XX
XX PD 20-FEB-2003.
XX

```



PF 01-AUG-2002; 2002WO-EP008671.  
 XX 03-AUG-2001; 2001DE-01038092.  
 XX (JERI-) JERINI AG.  
 XX Osterkamp F, Schutkowski M;  
 XX WPI; 2003-381387/36.  
 XX Immobilizing compounds on surfaces useful e.g. for producing peptide  
 PT arrays for determining enzyme specificity, based on formation of a  
 PT nitrene connecting group.  
 XX Disclosure; Page 2; 69pp; German.  
 XX This invention describes a novel method for immobilizing compounds on a  
 CC surface where each component carries a reactive group. During the  
 CC reaction a nitrene group is formed and the surface is planar. The nitrene  
 CC group is formed from reactive groups on the surface and the immobilizing  
 CC compounds, especially where the groups are hydroxylamino, preferably N-  
 CC substituted, aldehyde, ketone, acetal, ketal, thioacetal, thiocarbonyl or  
 CC thioacetal. The reaction is chemoselective, results in targeted  
 CC immobilization and may be promoted by microwave irradiation. The products  
 CC of the invention are used for localized immobilization of a wide variety  
 CC of biological compounds, especially to form arrays (biochips) of  
 CC peptides for determining substrate specificity of enzymes or for  
 CC purification, by selection of compounds that react with formation of a  
 CC nitrene group. When used to purify peptides, the method provides reliable  
 CC separation of by-products, particularly interrupted sequences.  
 CC Immobilization by nitrene formation can be performed selectively in  
 CC presence of other reactive groups, also in aqueous medium at pH 3-4.5,  
 CC i.e. where amino groups will not be reactive. The nitrene group allows  
 CC spectroscopic monitoring and immobilized compounds are easily released  
 CC (nitrenes are labile to dipolarophiles). This sequence represents a cell  
 CC adhesive peptide used to illustrate the method of the invention.  
 XX  
 XX Sequence 4 AA;  
 SQ  
 Query Match 100.0%; Score 26; DB 7; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 Db 1 RGDC 4  
 RESULT 14  
 AAR11587  
 ID AAR11587 standard; protein; 5 AA.  
 AC AAR11587;  
 XX  
 XX 12-JUN-1991 (first entry)  
 DT Fibrinogen receptor antagonist #4.  
 DE Fibrinogen receptor antagonist; thrombosis; fibrinogen;  
 XX fibrinogen receptor antagonist; thrombosis; fibrinogen;  
 KW fibrinogen receptor antagonist; thrombosis; fibrinogen;  
 XX fibrinogen receptor antagonist; thrombosis; fibrinogen;  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1. .5  
 FT Modified-site 2  
 FT Modified-site /label= Arg(phenyl)  
 XX EP422938-A.  
 PN 17-APR-1991.  
 PD 11-OCT-1990; 90EP-00311151.  
 XX

XX 13-OCT-1989; 89US-00421224.  
 XX (MERI) MERCK & CO INC.  
 XX Nutt RF, Brady ST, Veber KF, Duggan MF;  
 XX WPI; 1991-111423/16.  
 XX Polypeptide fibrinogen receptor antagonists - used to prevent thrombosis,  
 PT e.g. during cardiovascular surgery.  
 XX Disclosure; Page 7; 16pp; English.  
 XX Cys at position 1 carries an acetyl group. The peptide inhibits binding  
 CC of fibrinogen to the platelet membrane glycoprotein complex IIb/IIIa  
 CC receptor. It may be used to prevent post-operative thrombosis.  
 CC thromboembolism and reocclusion, platelet adhesion in extracorporeal  
 CC blood circulation systems and to prevent myocardial infarction. See also  
 CC AAR11584-6 and AAR11588-R11594  
 XX  
 XX Sequence 5 AA;  
 SQ  
 Query Match 100.0%; Score 26; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 Db 2 RGDC 5  
 RESULT 15  
 AAR10415  
 ID AAR10415 standard; protein; 5 AA.  
 XX  
 XX AAR10415;  
 XX  
 XX 10-APR-1991 (first entry)  
 DT Fibrinogen receptor antagonising peptide (III).  
 DE Fibrinogen receptor antagonist; platelet aggregation; thrombosis;  
 XX Fibrinogen receptor antagonist; platelet aggregation; thrombosis;  
 KW myocardial infarction.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1. .5  
 FT Modified-site 1. .1  
 FT Modified-site /label= N-MethylCys  
 XX EP410537-A.  
 PN 30-JAN-1991.  
 PD 23-JUL-1990; 90EP-00202015.  
 DE 28-JUL-1989; 89US-00386534.  
 XX (MERI) MERCK & CO INC.  
 XX Nutt RF, Brady SF, Veber DF;  
 XX WPI; 1991-030928/05.  
 XX New fibrinogen receptor antagonising peptide cpds. - used to inhibit  
 PT blood platelet aggregation during surgery on peripheral arteries and in  
 PT cardiovascular surgery.  
 XX Claim 5; Page 10; 10pp; English.  
 XX To residue N-MethylCys1 is attached Acetyl and to residue Cys5 is

CC attached OH. The peptide is prepd. using standard methods of solid phase  
 CC synthesis and is one of 17 pref. examples of a highly generic formula.  
 CC The peptide is used to inhibit fibrinogen-induced platelet aggregation  
 CC and is partic. useful because it does not significantly deplete the  
 CC platelet count. It has a relatively short duration of activity and is  
 CC thus useful where prevention of platelet aggregation over a short period  
 CC of time is desirable, e.g. in surgery of peripheral arteries and  
 CC cardiovascular surgery. It is also useful in prevention of platelet  
 CC thrombosis, thromboembolism and reocclusion either during and after  
 CC thrombolytic therapy, or after angioplasty of coronary and other arteries  
 CC and after coronary artery by-pass procedures, or to prevent myocardial  
 CC infarction. It may be combined with thrombolytic agents, e.g. plasminogen  
 CC activators or streptokinase, or anticoagulants. See also AAR10413-18 and  
 CC EP-410537, EP-410533, EP-410540, EP-410541, and EP-410767  
 XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 26; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 Db 2 RGDC 5

Search completed: April 16, 2004, 07:20:46  
 Job time : 45.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:55:57 ; Search time 43.7 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588b-2

Perfect score: 21

Sequence: 1 RGDV 4

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Geneseq\_29Jan04.\*
- 2: Geneseq\_1980s.\*
- 3: Geneseq\_1990s.\*
- 4: Geneseq\_2000s.\*
- 5: Geneseq\_2001s.\*
- 6: Geneseq\_2002s.\*
- 7: Geneseq\_2003as.\*
- 8: Geneseq\_2003bs.\*
- 9: Geneseq\_2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	2	AAR04609 Antiviral
2	21	100.0	4	2	AAR11751 Platelet
3	21	100.0	4	2	AAR14686 Hyaluron
4	21	100.0	4	2	AAR42558 Tri-/tetr
5	21	100.0	4	2	AAR56655 Platelet
6	21	100.0	4	2	AAR77843 RGDV pept
7	21	100.0	4	2	AAR62947 RGD contg
8	21	100.0	4	2	AAR22414 Vitroact
9	21	100.0	4	3	AAR80469 Cell adhe
10	21	100.0	4	3	AAR01570 Cell bind
11	21	100.0	4	4	AAR81475 Synthetic
12	21	100.0	4	4	AAR91951 Fibronect
13	21	100.0	4	4	AAR86863 Transport
14	21	100.0	4	5	AAR28386 Fibronect
15	21	100.0	4	5	AAR85698 Extra cel
16	21	100.0	5	1	AAR71157 Peptide w
17	21	100.0	5	2	AAR24514 Platelet
18	21	100.0	5	2	AAR25451 GPIIb/III
19	21	100.0	5	3	AAR54941 Peptide 1
20	21	100.0	5	4	AAR72588 Thrombin
21	21	100.0	5	5	ABO9883 Amino aci
22	21	100.0	6	1	AAR93735 Fibronoge
23	21	100.0	6	2	AAR39601 Leukocyte
24	21	100.0	6	2	AAR42563 Tri-/tetr
25	21	100.0	6	2	AAR52106 Targettin

26	21	100.0	6	2	ADE25450	GPIIb/III
27	21	100.0	6	3	AAY54940	Peptide 1
28	21	100.0	6	4	AAB81477	Synthetic
29	21	100.0	7	1	AAP92278	Fibrinoge
30	21	100.0	7	2	AAR24071	Cell-to-c
31	21	100.0	7	2	AAR48546	Integrin
32	21	100.0	7	2	AAR48522	Integrin
33	21	100.0	7	4	AAB75078	Nucleotid
34	21	100.0	7	4	AAB75078	Nucleotid
35	21	100.0	8	2	AAR69307	Gp. IIB/II
36	21	100.0	8	2	AAR48506	Integrin
37	21	100.0	8	2	AAR48526	Integrin
38	21	100.0	8	2	AAR48521	Integrin
39	21	100.0	8	2	AAR50590	GPIIb/III
40	21	100.0	8	2	ADE25487	Tc-99m la
41	21	100.0	8	3	AAY54973	Peptide 1
42	21	100.0	8	3	AAY95459	GPIIb/III
43	21	100.0	8	5	AAR03235	Kidins220
44	21	100.0	8	6	ABG72767	Human FNF
45	21	100.0	8	6	ABG72765	Human FNF

## ALIGNMENTS

RESULT 1  
AAR04609  
ID AAR04609 standard; protein; 4 AA.

XX AC AAR04609;  
XX DT 25-MAR-2003 (revised)  
XX DT 05-SEP-1990 (first entry)  
XX DE Antiviral agent.  
XX KW Antiviral; M2; poliovirus; polio; hepatitis.  
XX OS Synthetic.

XX JP02078631-A.

XX PD 19-MAR-1990.

XX PF 14-SEP-1988; 88JP-00228843.

XX PR 14-SEP-1988; 88JP-00228843.

XX PA (NIHA) NIPPON MINING CO.

XX WPI; 1990-129060/17.

PT Antiviral agent contg. tripeptide (unit) - of basic aminoacid, then alanine, glycine or sarcosine, and acidic aminoacid, effective against virus with protein-terminated DNA or RNA.  
PS Disclosure; Page ?; 4pp; Japanese.  
XX Peptide is effective against inhibiting propagation of DNA or RNA bonded, protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

Oy 1 RGDV 4

Db 1 RGDV 4

## RESULT 2

AAR11751  
ID AAR11751 standard; protein; 4 AA.

XX AC AAR11751;

XX 25-MAR-2003 (revised)

DT 04-JUL-1991 (first entry)

XX XX

DE Platelet aggregation inhibitor peptide.

XX XX

XX Platelet aggregation; thrombus; myocardial infarction; stroke;

KW peripheral arterial disease; blood clotting; thrombosis.

XX XX

OS Synthetic.

XX XX

PH Key Location/Qualifiers

FT Modified-site 2..2

FT /label= MeGly

XX XX

PN W09104746-A.

XX XX

PD 18-APR-1991.

XX XX

PF 23-SEP-1989; 89US-00415006.

XX XX

PR 29-SEP-1989; 89US-00415006.

XX XX

PR 07-JUN-1990; 90US-00534385.

XX XX

PA (RORE ) RORER INT OVERSEAS INC.

XX XX

PI Klein SI, Molino BF, Czekaj X, Gardner C, Pelletier J;

XX XX

DR WPI; 1991-132636/18.

XX XX

PT New peptide(s) and pseudo-peptide(s) - inhibit aggregation and thrombus

PT formation for treatment of myocardial infarction, stroke and peripheral

PT arterial disease.

XX XX

ES Claim 32; Page 42; 4pp; English.

XX XX

CC This peptide inhibits platelet aggregation and thrombus formation in

CC mammalian blood. It does this by inhibiting fibrinogen binding to

CC activated platelets and other adhesive glycoproteins involved in the

CC aggregation process. It is useful in a pharmaceutical compn. for the

CC prevention and treatment of thrombosis associated with certain disease

CC states such as myocardial infarction, stroke, peri- pheral arterial

CC disease and disseminated intravascular coagulation. Additionally it may

CC be useful for treating diseases associated with abnormal cell growth. See

CC also AAR11750 and AAR11752. (Updated on 25-MAR-2003 to correct PA field.)

XX XX

SQ Sequence 4 AA;

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 4;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4

DB 1 RGDV 4

## RESULT 3

AAR14686

ID AAR14686 standard; peptide; 4 AA.

XX XX

AC AAR14686;

XX XX

DT 25-MAR-2003 (revised)

DT 22-JAN-1992 (first entry)

XX XX

DE Hyaluronic acid-modified endothelial cell adhesion domain.

XX XX

KW GAG; reducing terminal residue-lactonised glycosaminoglycan.

XX OS Synthetic.

XX XX

PH Key Location/Qualifiers

FT Modified-site 1

FT /label= OTHER

FT /note= "see comments"

XX XX

PN EP454898-A.

XX XX

PD 06-NOV-1991.

XX XX

DP 12-OCT-1990; 90EP-00119607.

XX XX

PR 30-MAR-1990; 90JP-000811163.

XX XX

PA (SEGK ) SEIKAGAKU CORP.

PA (SEGK ) SEIKAGAKU KOGYO KK.

XX XX

PI Sakurai K, Miyazaki K;

XX XX

DR WPI; 1991-328878/45.

XX XX

PT New glycosaminoglycan-modified proteins - used to treat cancer,

PT inflammation and hereditary enzyme deficiency.

XX XX

PS Example 3; Page 22; 35pp; English.

XX XX

CC The N-terminus of this peptide is covalently joined to hyaluronic acid  
CC (hyaluronic acid content = 96.7 per cent). The amino group of the peptide  
CC is bonded to an aldehyde group which has been formed by reducing and  
CC partially oxidising a reducing terminal sugar moiety of a  
CC glycosaminoglycan. The C-terminal Valine is amidated. (Updated on 25-MAR-

CC 2003 to correct PA field.)

XX XX

SQ Sequence 4 AA;

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 4;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4

DB 1 RGDV 4

## RESULT 4

AAR42558

ID AAR42558 standard; peptide; 4 AA.

XX AC AAR42558;

XX XX

DT 04-NOV-1994 (first entry)

XX XX

DB Tri-/tetra-cyclic compound linked peptide #2.

XX XX

KW Tricyclic; tetracyclic; medicament; manufacture; determination;

KW blood platelet thrombi; fibrinogen antagonists.

XX OS Synthetic.

XX XX

PN AC9344935-A.

XX XX

PD 10-MAR-1994.

XX XX

PF 26-AUG-1993; 93AU-00044935.

XX XX

PR 31-AUG-1992; 92EH-00002725.

XX XX

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX XX

PI Banwarth W, Gerber R, Grieder A, Knierzinger A, Muller K;

PI Obrecht D, Trzeclak A;  
 XX WPI; 1994-118783/15.  
 XX New tri- and tetra-cyclic cpds., partic. with peptide chains - used e.g.  
 PT for determ. of biologically active peptide sequences or as fibrinogen  
 PT antagonists.  
 XX  
 XX C-aim 9; Page 120; 131pp; English.  
 XX  
 CC The sequences given in AAR42557-68 are peptides which may be linked to  
 CC the tri- and tetra-cyclic compounds of the invention. These compounds may  
 CC be used as medicaments or for the manufacture of medicaments. They may  
 CC esp. be used for the determination of biologically active peptide  
 CC sequences. Compounds containing the sequence RGDV or RGDV can esp. be  
 CC used for preventing the formation of blood platelet thrombi, as  
 CC fibrinogen antagonists. Claimed examples of this sequence include 4,5-  
 CC Cyclo[-acetyl-L-arginyl- glycy-L-aspartyl-L-valylaminomethyl-]-3,6-  
 CC dimethoxy-9,9-dimethyl- xanthene, 10-Methyl-4,6-cyclo[-acetyl-L-  
 CC arginyl-L-aspartyl- L-valylaminomethyl-]-phenothiazine, 4,6-Cyclo[-acetyl-L-  
 CC arginyl-glycyl- L-aspartyl-L-valylaminomethyl-]-3,7-dimethoxy-10-ethyl-10H-  
 CC dibenz[e,j][1,4]oxazine, 10-Hexyl-3,7-dimethoxy-4,6-cyclo[-acetyl-L-  
 CC arginyl- glycy-L-aspartyl-L-valylaminomethyl-]-phenothiazine, and 3,7-  
 CC Dimethoxy- 10-methyl-4,6-cyclo[-acetyl-L-arginyl-glycyl-L-aspartyl-L-  
 CC valylamino- methyl-]-phenothiazine  
 XX  
 XX Sequence 4 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 5  
 AAR56655  
 ID AAR56655 standard; protein; 4 AA.  
 XX  
 AC AAR56655;  
 DT 25-MAR-2003 (revised)  
 DT 14-MAR-1995 (first entry)  
 XX  
 DE Platelet aggregation inhibition peptide #2.  
 XX  
 KW Aspartic acid; inhibition; platelet aggregation; fibrinogen; bonding;  
 KW adhesion; glycoprotein; blood clotting; abnormal thrombus formation;  
 KW thrombosis; myocardial infarction; stroke; peripheral arterial disease.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 4  
 FT /note= "N-methyl-Val"  
 PN US5332726-A.  
 XX  
 PD 26-JUL-1994.  
 XX  
 XX 30-MAR-1992; 92US-00859779.  
 XX  
 XX 29-SEP-1989; 89US-00415006.  
 PR 04-JAN-1990; 90US-00460777.  
 PR 07-JUN-1990; 90US-00534385.  
 PR 28-MAR-1991; 91US-00677006.  
 XX  
 XX (RHON ) RHONE POULENC RORER PHARM INC.  
 PA  
 XX Becker MR, Klein SI, Gardner C, Czepak M, Dener JM, Molino BF;  
 XX

DR WPI; 1994-241043/29.  
 XX  
 PT New aspartic acid deriv. or analogue (pseudo)-peptide cpds. - are used as  
 PT platelet aggregation and thrombus formation inhibitors for treatment of  
 PT thrombosis.  
 XX  
 XX Example 5; Col 14; 24pp; English.  
 XX  
 CC The sequences given in AAR56654-56 are aspartic acid derivative peptides  
 CC which inhibit platelet aggregation by inhibiting fibrinogen bonding to  
 CC activated platelets and other adhesive glycoproteins involved in  
 CC platelet aggregation and blood clotting. These peptides may also be used  
 CC for the treatment or prevention of abnormal thrombus formation and for  
 CC the prevention of thrombosis associated with diseases such  
 CC as myocardial infarction, stroke and peripheral arterial disease. These  
 CC peptides are pref. prepared by solid phase or solution phase peptide  
 CC synthesis. (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 6  
 AAR77843  
 ID AAR77843 standard; peptide; 4 AA.  
 XX  
 AC AAR77843;  
 DT 24-JAN-1996 (first entry)  
 XX  
 DE RGDV peptide - epitope for monoclonal antibody.  
 XX  
 KW RGDV peptide; monoclonal antibody; inhibitor; blood cell activation;  
 KW aggregation; adhesion; retard coagulation; blood sample; analysis;  
 KW serine protease inhibitor.  
 XX  
 OS Synthetic.  
 XX  
 XX WO9514788-Ai.  
 EN  
 PD 01-JUN-1995.  
 XX  
 XX 23-NOV-1994; 94WO-US013537.  
 PF  
 XX 24-NOV-1993; 93US-00157880.  
 PR  
 XX (UNIW ) UNIV WASHINGTON.  
 PA  
 XX Lyon ME, Henderson P, Malik S, Kenny MA, Lyon AW;  
 XX  
 XX WPI; 1995-206945/27.  
 DR  
 XX Analysis of whole blood samples - using a serine protease inhibitor or a  
 XX blood cell activation, aggregation or adhesion inhibitor to retard  
 XX coagulation.  
 XX  
 XX Claim 4; Page 20; 27pp; English.  
 ES  
 CC Monoclonal antibodies (MAb) to the RGDV peptide, act as blood cell  
 CC activation, aggregation and adhesion inhibitors. The method of the  
 CC invention is used to measure analytes from whole blood, plasma and blood  
 CC cells. The method comprises, selecting a sample, retarding coagulation of  
 CC the sample with a mixture comprising a serine protease inhibitor (SPI),  
 CC and pref. a blood cell activation, aggregation and adhesion inhibitor  
 CC selected from annexins, MAb as above, triflavin and alpha, alpha-bis(3-  
 CC (N,N-diethylcarbonyl))-p-xylene dihydrobromide. The SPI is free of

CC interfering effects on detection of divalent cations by ion selective  
 CC electrodes, and is readily standardised for anti-coagulant dosage  
 CC response. Multiple analytes in a single coagulation retarded sample are  
 CC measured using analytical methods with accuracy dependent upon divalent  
 CC cation concn

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 ||||  
 Db 1 RGDV 4

RESULT 7

AAR62947  
 ID AAR62947 standard; peptide; 4 AA.

XX AC AAR62947;

XX DT 15-AUG-1995 (first entry)

XX RGD contg. peptide is a constituent of a cancer metastasis inhibitor.

XX RGD peptide; Arg-Gly-Asp; cancer metastasis inhibitor;

XX high cell adhesion; fibronectin.

XX Synthetic.

XX JP06306096-A.

XX 01-NOV-1994.

XX 31-JAN-1994; 94JP-00009893.

XX 26-FEB-1993; 93JP-00038677.

XX (FUJIF ) FUJIFILM CO LTD.

XX WPI; 1995-019276/03.

XX Peptide deriv. comprising several copies of the sequence Arg-Gly-Asp -  
 PT covalently bonded to organic molecule, useful as cancer metastasis  
 PT inhibitor.

XX Disclosure; Page 3; 13pp; Japanese.

XX AAR62945-50 are RGD contg. peptides that form constituents of an organic  
 CC mol. with a defined copy number of one of the peptides. The organic mol.  
 CC contg. the RGD peptide is a cancer metastasis inhibitor. The deriv. shows  
 CC substantially no toxicity and has a higher cell adhesion than the core  
 CC sequence of fibronectin

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 ||||  
 Db 1 RGDV 4

RESULT 8

AAY22414  
 ID AAY22414 standard; peptide; 4 AA.

XX AC AAY22414;

DT 28-SEP-1999 (first entry)  
 XX Vitronectin peptide fragment.  
 DE  
 XX BAL; biodegradable; biocompatible article; biologically active ligand;  
 KW tissue engineering template; tissue regeneration; nerve regeneration;  
 KW endothelial cell growth; vasculogenesis; wound repair; therapy;  
 KW vitronectin.  
 XX Synthetic.  
 XX WO9936107-A1.  
 XX 22-JUL-1999.  
 XX 20-JAN-1999; 99WO-GS300192.  
 XX 20-JAN-1998; 98GB-00001061.  
 XX (UNNO-) UNIV NOTTINGHAM.  
 XX (NASI ) MASSACHUSETTS INST TECHNOLOGY.  
 XX Shakesheff K, Patel N, Cannizzaro SM, Langer RS;  
 WPI; 1993-458396/38.  
 XX New biodegradable and biocompatible articles having a patterned array of  
 PT attached ligands, used for tissue regeneration or tissue engineering.  
 XX Disclosure; Page 17; 81pp; English.  
 CC This sequence represents a fragment of vitronectin, and can be used in  
 CC the article of the invention. The article is a biodegradable and  
 CC biocompatible article that comprises a surface and a biologically active  
 CC ligand (BAL) (such as this sequence) provided on the surface in a  
 CC spatially controlled pattern, where the BAL is attached to the surface by  
 CC a specific molecular interaction. The articles may be used as tissue  
 CC engineering templates in vivo or in vitro. They can be used in the  
 CC manufacture of a medicament or device for the treatment of a patient in  
 CC need of tissue regeneration or of a tissue engineering template, e.g.  
 CC nerve regeneration, endothelial cell growth, vasculogenesis or wound  
 CC repair. The methods can provide for the generation of micron-scale  
 CC patterns of biologically active ligands on biodegradable and  
 CC biocompatible article surfaces. A very wide range of pattern designs may  
 CC be formed using this technology  
 XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 ||||  
 Db 1 RGDV 4

RESULT 9

AAY80469  
 ID AAY80469 standard; peptide; 4 AA.

XX AC AAY80469;

XX 06-JUN-2000 (first entry)

XX Cell adhesion peptide #4.

XX Bone regenerative; osteopathic; osseous tissue; reconstruction;  
 KW scaffold matrix; bone formation promoter; bone resorption inhibitor;  
 KW cell adhesion; osteoblast; osteoclast; bone defect; fracture.  
 XX Synthetic.

PN WO200004941-A1.  
 XX 03-FEB-2000.  
 XX 22-JUL-1999; 99WO-US016800.  
 XX 24-JUL-1998; 98US-00122348.  
 XX (PHAR-) PHARMACAL BIOTECHNOLOGIES INC.  
 XX Budny JA;  
 XX WPI; 2000-135084/17.  
 DR System for reconstructing osseous tissue, useful e.g. for treating  
 XX fractures, comprises scaffold containing promoter of bone formation and  
 XX inhibitor of bone resorption.  
 XX Claim 14; Page 31; 44pp; English.  
 XX The invention relates to a novel system for reconstruction of osseous  
 XX tissue comprising a scaffold carrying a compound (I) that promotes bone  
 XX formation and a component that decreases bone resorption (II). (I)  
 XX induces migration and adhesion of osteoblasts and osteoclasts and (II)  
 XX inhibits proteolysis (specifically by plasmin) of extracellular matrix.  
 XX (I) is preferably selected from: selectin or selectin binding fragments,  
 XX proteins and peptides that facilitate cell adhesion, plasminogen  
 XX activator inhibitors, protease inhibitors and metalloprotease inhibitors.  
 XX The peptides AAY80466-Y80492 are claimed examples of cell adhesion  
 XX peptides used in the system of the invention. The system is used to  
 XX replace, remodel or correct bone defects, e.g. fractures, fissures or  
 XX bone mass loss. Incorporation of (I) into the scaffold results in rapid  
 XX seeding by osteoblasts and the development of an organic matrix, i.e. the  
 XX preformed scaffold replaces the rate-determining step of extracellular  
 XX matrix formation. The scaffold can be designed to have a predetermined  
 XX resorption/degradation rate, and may include regulatory compounds for  
 XX specific cell types  
 XX SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 3; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 Db 1 RGDV 4  
 ||||  
 1 RGDV 4  
 RESULT 10  
 AAB01570  
 ID AAB01570 standard; peptide; 4 AA.  
 XX AC AAB01570;  
 XX 08-NOV-2000 (first entry)  
 DT Cell binding domain of vitronectin.  
 XX Polymer; biomaterial; conjugate; hydrogel; drug delivery; adhesive;  
 KW sealant; tissue engineering; wound healing; scaffold; cell transplant;  
 KW adhesion prevention; cell migration; collagenase; plasmin; elastase.  
 XX Synthetic.  
 OS WO200004808-A1.  
 XX 03-AUG-2000.  
 XX 01-FEB-2000; 2000WO-US002608.  
 XX 01-FEB-1999; 99US-0118093P.  
 XX

(HUBB/) HUBBELL J A.  
 Hubbell JA, Elbert D, Lutolf M, Pratt A, Schoenmakers R;  
 Tirelli N, Vernon B;  
 WPI; 2000-524289/47.  
 Producing polymeric biomaterials by polymerizing two or more precursor  
 components (e.g. polymer, protein or peptide) of the biomaterial, useful  
 for delivering therapeutic molecules to a subject and as adhesives or  
 sealants.  
 Disclosure; Page 52; 119pp; English.  
 A method of making polymeric biomaterials is described comprising  
 combining two or more precursor components (e.g. polymer, protein or  
 peptide) of the biomaterial under conditions that allow polymerisation of  
 the two components. Polymerisation occurs through self selective reaction  
 between a strong nucleophile and a conjugated unsaturated bond or a  
 conjugated unsaturated group, by nucleophilic addition. The polymeric  
 hydrogels can be used in a variety of applications. They can be used to  
 deliver therapeutic molecules to a subject, as adhesives or sealants  
 (e.g. sealing air leaks on the lung), as tissue engineering and wound  
 healing scaffolds, and as cell transplant devices. The biomaterials are  
 also useful for adhesion prevention to minimise unwanted operative or  
 post-traumatic adhesions. A variety of adhesion-promoting peptides have  
 been identified as being the active domains of adhesion-promoting  
 proteins such as fibronectin, vitronectin, laminin, collagen, von  
 Willebrand factor osteonectin etc. These peptides can be incorporated  
 into the biomaterial when they are designed with a strong nucleophile in  
 the peptide chain such as cysteine. These peptides are potentially useful  
 in controlling a variety of cellular reactions such as cell attachment,  
 migration and overgrowth on a material surface when the material is non  
 bio-degradable or slowly degradable, and cell migration through a  
 material when that material is biodegradable. The peptides are also  
 useful in the induction of particular cellular phenotypes  
 XX SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 3; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 Db 1 RGDV 4  
 ||||  
 1 RGDV 4  
 RESULT 11  
 AAB81475  
 ID AAB81475 standard; peptide; 4 AA.  
 XX AC AAB81475;  
 XX 18-JUN-2001 (first entry)  
 DT Synthetic peptide, SEQ ID NO: 21.  
 DE Peptidomimetic; beta-hairpin loop mimetic; binding motif identification;  
 KW drug design; protein target identification.  
 XX Synthetic.  
 OS WO200116161-A1.  
 XX 08-MAR-2001.  
 XX 30-AUG-1999; 99WO-EP006369.  
 XX 30-AUG-1999; 99WO-EP006369.  
 XX (POLY-) POLYPHOS AG.  
 XX

PI Robinson CA, Obrecht D;  
 XX WPI; 2001-273332/28.  
 XX  
 XX Method for manufacturing template-fixed beta-hairpin loop mimetics,  
 PT useful for designing small peptidomimetic drug candidates, involves  
 PT process based on mixed solid and solution phase synthetic strategy.  
 XX  
 XX Claim 7; Page 77; 83pp; English.  
 XX  
 CC The present sequence is part of a template-fixed, beta-hairpin loop  
 CC peptidomimetic which may be manufactured by the process disclosed in this  
 CC invention. The process is based on a mixed solid and solution phase  
 CC synthetic strategy. The method is useful for determining key amino acids  
 CC and motifs important for binding large surface and flat protein  
 CC interfaces in their sequential and/or spatial arrangement. This  
 CC information can ultimately be used for the design of small peptidomimetic  
 CC drug candidates. The peptidomimetics may be used to probe large surface  
 CC protein-protein interactions and to find protein targets  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 4; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 12  
 AAB91951  
 ID AAB91951 standard; peptide; 4 AA.  
 AC AAB91951;  
 XX  
 DT 22-JUN-2001 (first entry)  
 DE Fibronectin fragment and fibrin related peptide SEQ ID NO:1127.  
 XX  
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimide; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FN WO2000069303-A2.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 17-MAY-2000; 2000WO-US013576.  
 XX  
 PR 17-MAY-1999; 99US-0134406P.  
 PR 10-SEP-1999; 99US-0153406P.  
 PR 15-OCT-1999; 99US-0159783P.  
 XX  
 PA (CONJ-) CONJUCHEM INC.  
 XX  
 Z- Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
 XX  
 DR WPI; 2001-112059/-2.  
 XX  
 PT Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity.  
 XX  
 PS Disclosure; Page 564; 733pp; English.  
 XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimide and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently

CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity in  
 CC vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 2; DB 4; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 13  
 AAB86863  
 ID AAB86863 standard; peptide; 4 AA.  
 AC AAB86863;  
 XX  
 DT 28-NOV-2001 (first entry)  
 DE Transport molecule/ligand binding-associated peptide #9.  
 XX  
 KW Transport molecule; ligand; cancer treatment; autoimmune disease;  
 KW inflammation; infection.  
 XX  
 OS Synthetic.  
 XX  
 FN WO200168142-A1.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 13-MAR-2001; 2001WO-EF002833.  
 XX  
 PR 13-MAR-2000; 2000DE-01012120.  
 XX  
 PA (KTBT-) KTB TUMORFORSCHUNGS GMBH.  
 XX  
 PI Kratz F;  
 XX  
 DR WPI; 2001-589998/66.  
 XX  
 PT New ligand, comprising therapeutic or diagnostic agent bonded non-  
 PT covalently with substance having high affinity to transport molecule.  
 XX  
 PS Disclosure; Page 39; 74pp; German.  
 XX  
 CC This invention describes novel ligands which bind to transport molecules,  
 CC comprising a therapeutic and/or diagnostic agent (A) non-covalently  
 CC bonded via a linkage cleavable in vivo depending on pH and/or  
 CC enzymatically with a substance (B) having an association constant KA to a  
 CC transport molecule of above 10<sup>3</sup> M<sup>-1</sup>, is new. The medicaments are  
 CC especially useful for the treatment of cancers, autoimmune diseases,  
 CC acute and chronic inflammation and infections caused by viruses or  
 CC microorganisms. The diagnostic kits are useful for the detection of these  
 CC illnesses and for the detection of the transport molecule and/or its  
 CC distribution in vivo. The ligands have excellent solubility in the medium  
 CC at the site of action and are easy and inexpensive to convert into  
 CC adducts, as the interaction with the transport material is physical.  
 CC AAB86843-AAB86920 represent peptides used to illustrate the method of the  
 CC invention



XX SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 4; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 14  
 AA228386  
 ID AAE228386 standard; peptide; 4 AA.  
 XX AC  
 AC AAE228386;  
 DT 27-DEC-2002 (first entry)  
 XX DE  
 DE Fibronectin 2 RGD cell binding region #2.  
 KW Tat region; nucleic acid-binding group; cell transfection system;  
 KW gene therapy; cancer; fibronectin 2.  
 XX CS Unidentified.  
 XX FN US6376248-B1.  
 XX PD 23-APR-2002.  
 XX PF 16-MAR-1998; 98US-0003978C.  
 XX PR 14-MAR-1997; 97US-0081823C.  
 XX PA (LIFE-) LIFE TECHNOLOGIES INC.  
 XX PI Hawley-Nelson P, Lan J, Shih P, Jesse JA, Schifferli KP;  
 FI Gebeyehu G, Ciccarone VC, Evans KL;  
 XX WI; 2002-680647/73.  
 XX PT New peptide comprising Tat sequence linked to nucleic acid-binding group,  
 PT useful, e.g. in gene therapy, for improving cell-transfection efficiency.  
 XX Example 1; Col 61; 108pp; English.  
 XX The invention relates to a peptide comprising Tat sequence linked to  
 CC nucleic acid-binding group. Peptides of the invention are used as  
 CC components of a cell transfection system particularly for gene therapy  
 CC (especially of cancer). The present sequence is fibronectin 2 RGD cell  
 CC binding region. This peptide is used in the exemplification of the  
 CC invention  
 XX SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 5; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 DB 1 RGDV 4  
 RESULT 15  
 AAU85698  
 ID AAC85698 standard; peptide; 4 AA.  
 XX AC  
 AC AAU85698;  
 DT 29-JUG-2003 (revised)  
 DT 2-MAY-2002 (first entry)

XX Extra cellular matrix protein cell binding domain #3.  
 DE Biomaterial; vasotropic; anticoagulant; thrombolytic; vulnery;  
 XX infection; adhesion; thrombosis; restenosis; adhesive; sealant;  
 XX tissue engineering; wound healing scaffold; cell transplant device.  
 XX unidentified.  
 XX WO200192584-A1.  
 XX 06-DEC-2001.  
 XX 04-JUN-2001; 2001WO-US018101.  
 XX 02-JUN-2000; 200JUS-00586937.  
 XX (EIDG-) EIDGENESSISCHE TECH HOCHSCHULE ZUERICH.  
 XX (UYZU-) UNIV ZURICH.  
 XX Hubbel JA, Elbert D, Schoenmakers R;  
 XX WI; 2002-205802/26.  
 XX PT New biomaterial useful for medical treatment comprises an active or a  
 PT binding group and has an ester or amide bond onto the active or binding  
 PT group.  
 XX Disclosure; Page 90; 222pp; English.  
 XX The invention relates to a biomaterial comprising an active or a binding  
 CC group and has an ester or amide bond onto the active or binding group.  
 CC Also included is a biomaterial (II) formed from the cross-linking of at  
 CC least two precursor components of formula D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-S-(CH<sub>2</sub>)<sub>2</sub>-  
 CC COX-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-(CH<sub>2</sub>)<sub>2</sub>-COX-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-U-P',  
 CC D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-S-U-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-L-S-CH<sub>2</sub>-CH<sub>2</sub>-COX-P', D-Y-  
 CC C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-S-L-S-U-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-L-S-CH<sub>2</sub>-CH<sub>2</sub>-COX-P', D-Y-  
 CC C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-L-S-U-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-S-L-NH-CH<sub>2</sub>-CH<sub>2</sub>-COX-P',  
 CC D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-S-L-NH-U-P', D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-L-NH-CH<sub>2</sub>-CH<sub>2</sub>-COX-  
 CC P'; or D-Y-C(O)-(CH<sub>2</sub>)<sub>2</sub>-n-NH-L-NH-U-P'. The half-life of the ester or  
 CC amide bond onto the active or binding group is 1 hour - 1 year in an  
 CC aqueous solution at pH 7.4 and 37 plusC. Forming (II) involves (a)  
 CC attaching the active or binding compound to the linker molecule or  
 CC incorporating a nucleophilic amine or thiol into the active or binding  
 CC compound, (b) removing any thiol- or amine-protecting groups in the  
 CC linker (c) coupling a thiol, amine or alkene in the linker or  
 CC incorporated into the active or binding compound to the water soluble or  
 CC water swellable polymer containing conjugated unsaturated groups by a  
 CC conjugate addition reaction to form a precursor component and (d) cross-  
 CC linking the uncoupled conjugated unsaturated groups in at least one  
 CC precursor component. D = active or binding group; Y = O, NH or N; L =  
 CC linear or branched linker; X = O or N; P' = water-soluble or water-  
 CC swellable polymer containing conjugated unsaturated groups; U = product  
 CC of the addition of a nucleophile to an electrophile that is attached to  
 CC the water-soluble or water-swellable polymer; and n = 2 - 3. The  
 CC biomaterials are useful for treating or preventing disease, disorder or  
 CC infection in a mammal e.g. human, for preventing adhesions, thrombosis or  
 CC restenosis in a mammal and for delivering the active compound (preferably  
 CC protein or peptide) to a cell, tissue, organ, organ system or a body of a  
 CC mammal as adhesives or sealants, as tissue engineering and wound healing  
 CC scaffolds and as cell transplant devices. The bond of the biomaterial has  
 CC a half life of 1 hour - 1 year (preferably 1 day - 9 months, especially 2  
 CC days - 2 months) in an aqueous solution at pH 7.4 and 37plusoc. Thus the  
 CC biomaterial releases the therapeutic compound over a clinically relevant  
 CC time-frame. The present sequence is a peptide which acts as a  
 CC therapeutically active compound suitable for inclusion in the biomaterial  
 CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)  
 XX SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 5; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGV 4  
Db 1 RGV 4

Search completed: April 16, 2004, 07:20:48  
Job time : 45.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 12.9 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-2

Perfect score: 21

Sequence: 1 RGDV 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap:\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	1	US-08-128-225-1
2	21	100.0	4	1	US-08-658-130-13
3	21	100.0	4	2	US-08-668-871-2
4	21	100.0	4	2	US-08-669-683-2
5	21	100.0	4	4	US-09-039-780A-16
6	21	100.0	4	5	PCT-US96-08723-13
7	21	100.0	5	1	US-07-602-847C-34
8	21	100.0	5	1	US-08-497-134A-34
9	21	100.0	5	1	US-08-482-880-15
10	21	100.0	5	2	US-08-273-274-15
11	21	100.0	5	2	US-08-475-041-15
12	21	100.0	5	2	US-08-484-773-15
13	21	100.0	5	2	US-08-361-864-19
14	21	100.0	5	4	US-09-375-808-25
15	21	100.0	5	6	5196510-32
16	21	100.0	5	6	5384309-8
17	21	100.0	5	6	US-07-866-678-5
18	21	100.0	6	1	US-08-482-880-14
19	21	100.0	6	2	US-08-668-871-7
20	21	100.0	6	2	US-08-669-683-7
21	21	100.0	6	2	US-08-273-274-14
22	21	100.0	6	2	US-08-702-673-3
23	21	100.0	6	2	US-08-475-241-14
24	21	100.0	6	2	US-08-753-781-26
25	21	100.0	6	2	US-08-484-773-14
26	21	100.0	6	2	US-08-361-864-18
27	21	100.0	6	5	PCT-US91-08328-34

28	21	100.0	7	2	US-08-753-781-25	Sequence 25, Appl
29	21	100.0	7	5	PCT-US91-08328-5	Sequence 5, Appl
30	21	100.0	7	5	PCT-US91-08328-33	Sequence 33, Appl
31	21	100.0	8	2	US-08-335-832-15	Sequence 15, Appl
32	21	100.0	8	2	US-08-361-864-32	Sequence 32, Appl
33	21	100.0	8	3	US-09-141-127-9	Sequence 9, Appl
34	21	100.0	8	5	PCT-US91-08328-32	Sequence 32, Appl
35	21	100.0	8	6	5196510-38	Patent No. 5196510
36	21	100.0	9	1	US-08-482-880-17	Sequence 17, Appl
37	21	100.0	9	1	US-08-421-702A-6	Sequence 6, Appl
38	21	100.0	9	1	US-08-303-052A-6	Sequence 6, Appl
39	21	100.0	9	1	US-08-421-696A-6	Sequence 6, Appl
40	21	100.0	9	1	US-08-421-697A-6	Sequence 6, Appl
41	21	100.0	9	1	US-08-421-698A-6	Sequence 6, Appl
42	21	100.0	9	2	US-08-273-274-17	Sequence 17, Appl
43	21	100.0	9	2	US-08-421-695A-6	Sequence 6, Appl
44	21	100.0	9	2	US-08-475-041-17	Sequence 17, Appl
45	21	100.0	9	2	US-08-335-832-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-08-128-225-1  
; Sequence 1, Application US/08128225  
; Patent No. 5691203  
; GENERAL INFORMATION:  
; APPLICANT: KATSUEN, Susumu  
; APPLICANT: OHSHIMA, Kunihiko  
; APPLICANT: YAMAMOTO, Rychei  
; APPLICANT: NISHINO, Toyokazu  
; TITLE OF INVENTION: METHOD FOR CULTURING ANIMAL ADHESIVE  
; TITLE OF INVENTION: CELLS IN SERUM-FREE CONDITION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: WEGNER, CANTOR, MUELLER & PLAYER  
; STREET: 1233 20th Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-8218  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/128,225  
; FILING DATE: 29-SEP-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 141984/1992  
; FILING DATE: 14-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cantor, Herbert I.  
; REGISTRATION NUMBER: 24,392  
; REFERENCE/DOCKET NUMBER: P-500-24015  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-0400  
; TELEFAX: (202) 835-0605  
; TELEX: 440706 and 248394  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-128-225-1  
Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
|||  
Db 1 RGDV 4

## RESULT 2

US-08-658-130-13  
; Sequence 13, Application US/08658130  
; Patent No. 5736392  
; GENERAL INFORMATION:  
; APPLICANT: Hawley-Nelson, Pamela  
; APPLICANT: Lan, Jiaqing  
; APPLICANT: Shin, Posen  
; APPLICANT: Jesse, Joel A.  
; APPLICANT: Shifferli, Kevin P.  
; TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid  
; TITLE OF INVENTION: Transfections  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/658,130  
FILING DATE: 04-JUN-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/477,354  
FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Caruthers, Jennie M.  
REGISTRATION NUMBER: 34,464  
REFERENCE/DOCKET NUMBER: 32-95A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal

US-08-658-130-13

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
|||  
Db 1 RGDV 4

## RESULT 3

US-08-668-871-2  
; Sequence 2, Application US/08668871  
; Patent No. 5811389  
; GENERAL INFORMATION:  
; APPLICANT: Bannwarth, Wilhelm  
; APPLICANT: Gerber, Ferdinand

; APPLICANT: Grieder, Alfred  
; APPLICANT: Knierzinger, Andreas  
; APPLICANT: Mueller, Klaus  
; APPLICANT: Obrecht, Daniel  
; APPLICANT: Trzeciak, Arnold  
; TITLE OF INVENTION: TRI- AND TETRACYCLIC COMPOUNDS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: New Jersey  
; COUNTRY: United States of America  
; ZIP: 07110-1199

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/668,871  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/106,508  
FILING DATE:  
APPLICATION NUMBER: CH 2725/92  
FILING DATE: 31-AUG-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Kass, Alan P  
REGISTRATION NUMBER: 32142  
REFERENCE/DOCKET NUMBER: RAN 4781/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-4205  
TELEFAX: (201) 235-3530  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-668-871-2

Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
|||  
Db 1 RGDV 4

## RESULT 4

US-08-669-683-2  
; Sequence 2, Application US/08669683  
; Patent No. 5811548  
; GENERAL INFORMATION:

; APPLICANT: Bannwarth, Wilhelm  
; APPLICANT: Gerber, Ferdinand  
; APPLICANT: Grieder, Alfred  
; APPLICANT: Knierzinger, Andreas  
; APPLICANT: Mueller, Klaus  
; APPLICANT: Obrecht, Daniel  
; APPLICANT: Trzeciak, Arnold  
; TITLE OF INVENTION: TRI- AND TETRACYCLIC COMPOUNDS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: New Jersey  
; COUNTRY: United States of America  
; ZIP: 07110-1199

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,683
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/106,508
; FILING DATE:
; APPLICATION NUMBER: CH 2725/92
; FILING DATE: 31-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: RAN 4781/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-669-683-2

Query Match 100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
Db 1 RGDV 4
|||||

RESULT 5
US-09-039-780A-16
; Sequence 16, Application US/09039780A
; Patent No. 6376248
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NELSON, PAMELA
; LAN, JIANQING
; SEIH, POJEN
; JESSE, JOEL A.
; SCHIFFERLI, KEVIN P.
; GEBVEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,780A
; FILING DATE: 16-Mar-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
```

```
;
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-039-780A-16

Query Match 100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
Db 1 RGDV 4
|||||

RESULT 6
PCT-US96-08723-13
; Sequence 13, Application PC/TUS9608723
; GENERAL INFORMATION:
; APPLICANT: Life Technologies, Inc.
; TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid
; TITLE OF INVENTION: Transfections
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08723
; FILING DATE: 04-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 3295A WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080\
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; PCT-US96-08723-13

Query Match 100.0%; Score 21; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
|||||
```

Db 1 RGDV 4

## RESULT 7

US-07-602-847C-34  
; Sequence 34, Application US/07602847C  
; Patent No. 5227469  
; GENERAL INFORMATION:  
; APPLICANT: Lazarus, Robert A.,  
; APPLICANT: Seymour, Jana L.  
; TITLE OF INVENTION: NO. 5227469el Platelet Aggregation Inhibitors From The Leech  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/602,847C  
; FILING DATE: 19901026  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/479,829  
; FILING DATE: 14-FEB-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winter, Daryl B.  
; REGISTRATION NUMBER: 32,637  
; REFERENCE/DOCKET NUMBER: 667  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/266-1249  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
US-07-602-847C-34

Query Match 100.0%; Score 21; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4

DB 2 RGDV 5

## RESULT 8

US-08-497-134A-34  
; Sequence 34, Application US/08497134A  
; Patent No. 5667782  
; GENERAL INFORMATION:  
; APPLICANT: ROY, Polly  
; TITLE OF INVENTION: MULTIPLE PARTICULATE ANTIGEN DELIVERY  
; SYSTEM  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixoc & Vanderhye, P.C.  
; STREET: 1100 No. 5667782th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/497,134A  
; FILING DATE: 30-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mitchard, Leonard C.  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 1150-64  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-497-134A-34

Query Match 100.0%; Score 21; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4

DB 2 RGDV 5

## RESULT 9

US-08-482-880-15  
; Sequence 15, Application US/08482880  
; Patent No. 5736122  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T  
; APPLICANT: Lister-James, John  
; TITLE OF INVENTION: Tectonium-99m Labeled Peptides for  
; Thrombus Imaging  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, Ltd.  
; STREET: 10 South Wacker Drive Suite 3000  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,880  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5736122man, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 92,216-L  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 910-221-5317  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-482-880-15

Query Match 100.0%; Score 21; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 2 RGDV 5

## RESULT 10

US-08-273-274-15  
 ; Sequence 15, Application US/08273274  
 ; Patent No. 5849260  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dean, Richard T  
 ; APPLICANT: Lister-James, John  
 ; TITLE OF INVENTION: Technetium-99m Labeled Peptides for  
 ; TITLE OF INVENTION: Thrombus Imaging  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Allegretti & Witcoff, Ltd.  
 ; STREET: 10 South Wacker Drive Suite 3000  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/273,274  
 ; FILING DATE:  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/886,752  
 ; FILING DATE: 21-MAY-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5849260nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 92,216  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-715-1000  
 ; TELEFAX: 312-715-1234  
 ; TELEX: 910-221-5317  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-273-274-15

Query Match 100.0%; Score 21; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 2 RGDV 5

## RESULT 11

US-08-475-041-15  
 ; Sequence 15, Application US/08475041  
 ; Patent No. 5879658  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dean, Richard T  
 ; APPLICANT: Lister-James, John  
 ; TITLE OF INVENTION: Technetium-99m Labeled Peptides for  
 ; TITLE OF INVENTION: Thrombus Imaging  
 ; NUMBER OF SEQUENCES: 33

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, Ltd.  
 STREET: 10 South Wacker Drive Suite 3000  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/475,041  
 FILING DATE: 07-JUN-1995

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5879658nan, Kevin E  
 REGISTRATION NUMBER: 35,303  
 REFERENCE/DOCKET NUMBER: 92,216-M

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000  
 TELEFAX: 312-715-1234  
 TELEX: 910-221-5317  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-475-041-15

Query Match 100.0%; Score 21; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 2 RGDV 5

## RESULT 12

US-08-484-773-15  
 ; Sequence 15, Application US/08484773  
 ; Patent No. 5968476  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Deat, Richard T  
 ; APPLICANT: Lister-James, John  
 ; TITLE OF INVENTION: Technetium-99m Labeled Peptides for  
 ; TITLE OF INVENTION: Thrombus Imaging  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Allegretti, Ltd.  
 ; STREET: 10 South Wacker Drive Suite 3000  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484,773  
 FILING DATE: 07-JUN-1995

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5968476nan, Kevin E  
 REGISTRATION NUMBER: 35,303  
 REFERENCE/DOCKET NUMBER: 92,216-O  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-773-15

Query Match 100.0%; Score 21; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 2 RGDV 5

## RESULT 13

US-08-361-864-19  
Sequence 19, Application US/08361864  
Patent No. 5977064  
GENERAL INFORMATION:  
APPLICANT: Dean, Richard T  
APPLICANT: Lister-James, John  
TITLE OF INVENTION: Multimeric Polyvalent Antithrombotic  
TITLE OF INVENTION: Agents  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,864  
FILING DATE:

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/955,466A  
FILING DATE: 19921002

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5977064nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 92,668  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-361-864-19

Query Match 100.0%; Score 21; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 2 RGDV 5

RESULT 14  
US-09-375-808-25  
Sequence 25, Application US/09375808  
Patent No. 6544750  
GENERAL INFORMATION:  
APPLICANT: Schmaier, Alvin H.  
APPLICANT: Hasan, Ahmed A.K.

TITLE OF INVENTION: PEPTIDE ANALOGS AS SELECTIVE INHIBITORS OF THROMBIN  
TITLE OF INVENTION: ACTIVATION OF PROTEASE ACTIVATED RECEPTOR 1  
FILE REFERENCE: 32921-149520  
CURRENT APPLICATION NUMBER: US/09/375,808  
CURRENT FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 25  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: peptide  
US-09-375-808-25

Query Match 100.0%; Score 21; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 1 RGDV 4

## RESULT 15

5196510-32  
Patent No. 5196510  
APPLICANT: RODWELL, JOHN D.; MCKEARN, THOMAS J.; ALVAREZ,  
VERNON, L.; RADCLIFFE, ROBERT D.

TITLE OF INVENTION: MOLECULAR RECOGNITION UNITS

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/519,702  
FILING DATE: 07-MAY-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 291,730  
FILING DATE: 23-DEC-1988  
SEQ ID NO: 32:  
LENGTH: 5

Query Match 100.0%; Score 21; DB 6; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 1 RGDV 4

Search completed: April 16, 2004, 07:31:37  
Job time : 12.9 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 16.125 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-4  
Perfect score: 27  
Sequence: 1 GRGDS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 359414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 359414

Minimum DB seq length: 3  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgr2\_6/prodata/2/iaa/5A.COMB.pap.\*  
2: /cgr2\_6/prodata/2/iaa/5B.COMB.pap.\*  
3: /cgr2\_6/prodata/2/iaa/6A.COMB.pap.\*  
4: /cgr2\_6/prodata/2/iaa/6B.COMB.pap.\*  
5: /cgr2\_6/prodata/2/iaa/PCFUS.COMB.pap.\*  
6: /cgr2\_6/prodata/2/iaa/backfiles1.pap.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	1	US-07-780-790A-8
2	27	100.0	5	1	US-07-834-848-17
3	27	100.0	5	1	US-07-973-235A-16
4	27	100.0	5	1	US-08-251-027-15
5	27	100.0	5	1	US-08-178-482-12
6	27	100.0	5	1	US-08-259-569-7
7	27	100.0	5	1	US-08-240-711-9
8	27	100.0	5	1	US-08-173-115-1
9	27	100.0	5	1	US-08-457-753-9
10	27	100.0	5	1	US-08-482-106-18
11	27	100.0	5	1	US-08-482-107B-1
12	27	100.0	5	2	US-08-462-720-16
13	27	100.0	5	2	US-08-826-885-7
14	27	100.0	5	2	US-08-747-137-71
15	27	100.0	5	2	US-08-747-137-181
16	27	100.0	5	3	US-08-972-760-83
17	27	100.0	5	3	US-09-164-021-10
18	27	100.0	5	3	US-09-089-645A-83
19	27	100.0	5	3	US-09-480-332-8
20	27	100.0	5	3	US-09-677-714-1
21	27	100.0	5	3	US-08-476-134A-13
22	27	100.0	5	3	US-08-442-108B-27
23	27	100.0	5	3	US-09-423-347-4
24	27	100.0	5	4	US-09-503-998-83
25	27	100.0	5	4	US-03-197-770B-34
26	27	100.0	5	4	US-09-626-302-1
27	27	100.0	5	4	US-09-485-077A-7

28	27	100.0	5	4	US-09-638-282A-72	Sequence 72, Appl
29	27	100.0	5	4	US-09-134-253-5	Sequence 5, Appl
30	27	100.0	5	4	US-09-939-481-83	Sequence 83, Appl
31	27	100.0	5	4	US-09-548-697A-3	Sequence 3, Appl
32	27	100.0	5	4	US-09-096-749A-72	Sequence 72, Appl
33	27	100.0	5	4	US-08-918-189A-8	Sequence 8, Appl
34	27	100.0	5	6	5190920-8	Patent No. 5190920
35	27	100.0	5	6	5330911-11	Patent No. 5330911
36	27	100.0	5	6	5506208-9	Patent No. 5506208
37	27	100.0	6	1	US-07-720-222-46	Sequence 46, Appl
38	27	100.0	6	1	US-07-866-678-4	Sequence 4, Appl
39	27	100.0	6	1	US-07-780-790A-6	Sequence 6, Appl
40	27	100.0	6	1	US-07-932-200-11	Sequence 11, Appl
41	27	100.0	6	1	US-08-234-979-3	Sequence 3, Appl
42	27	100.0	6	1	US-08-251-027-13	Sequence 13, Appl
43	27	100.0	6	1	US-08-214-770-13	Sequence 13, Appl
44	27	100.0	6	1	US-08-212-186A-2	Sequence 2, Appl
45	27	100.0	6	1	US-08-169-743-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-07-780-790A-8  
; Sequence 8, Application US/07780790A  
; Patent No. 5298488  
; GENERAL INFORMATION:  
; APPLICANT: KOJIMA, Masayoshi  
; APPLICANT: KOMAZAWA, Hiroyuki  
; TITLE OF INVENTION: OM-CHITIN DERIVATIVES AND USE THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Sughrue, Mion, Zinn, Macpeak, & Seas  
; STREET: 2100 Pennsylvania Avenue  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: USA  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07780,790A  
; FILING DATE: 19911023  
; CLASSIFICATION: 530  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-780-790A-8

Query Match 100.0%; Score 27; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5

RESULT 2  
US-07-834-848-17  
; Sequence 17, Application US/07834848  
; Patent No. 5436221

```

; GENERAL INFORMATION:
; APPLICANT: KITAGUCHI, HIROSHI
; APPLICANT: KOMAZAWA, HIROYUKI
; APPLICANT: KAJIMA, MASAYOSHI
; APPLICANT: MORI, HIDETO
; APPLICANT: NISHIKAWA, NAOMYUKI
; APPLICANT: SATOH, HIDEAKI
; APPLICANT: OIKASA, ATSUSHI
; APPLICANT: ONO, MITSUNORI
; APPLICANT: AZUMA, ICHIRO
; APPLICANT: SAKI, IKUO
; TITLE OF INVENTION: PEPTIDE DERIVATIVES AND APPLICATION
; THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak, & Seas
; STREET: 2100 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/834,848
; FILING DATE: 19920213
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Biggart, Waddell A.
; REGISTRATION NUMBER: 24,861
; REFERENCE/DOCKET NUMBER: Q28480
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-834-848-17

Query Match 100.0%; Score 27; DB -: Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 1 GRGDS 5

RESULT 3
US-07-973-235A-16
; Sequence 16, Application US/07973235A
; Patent No. 5491130
; GENERAL INFORMATION:
; APPLICANT: David D. Roberts, et al.
; TITLE OF INVENTION: Peptide Inhibitors of Fibronectin and
Related Collagen-Binding Proteins
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, LeBlanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk

```

```

; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect, Version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,235A
; FILING DATE: 19921110
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert L. Price
; REGISTRATION NUMBER: 22,685
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)684-1111
; TELEFAX: (703)684-1124
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACIDS
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-973-235A-16

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 1 GRGDS 5

RESULT 4
US-08-251-027-15
; Sequence 15, Application US/08251027
; Patent No. 5519005
; GENERAL INFORMATION:
; APPLICANT: Lider, Ofer
; APPLICANT: Greenspoon, No. 5519005-
; APPLICANT: Hershkoviz, Rami
; APPLICANT: Alon, Ronen
; TITLE OF INVENTION: A METHOD OF INHIBITION OF CELLULAR AND
MOLECULAR LEVEL BIOLOGICAL INTERACTIONS UTILIZING
NON-PEPTIDIC SURROGATES OF THE ARG-GLY-ASP SEQUENCE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5519005thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,027
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 539-5050
; TELEFAX: (810) 539-5055
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

## US-08-251-027-15

Query Match 100.0%; Score 27; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 5

US-08-178-482-12  
 ; Sequence 12, Application US/08178482  
 ; Patent No. 5629294  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIZBERGA, GERE S  
 ; APPLICANT: RODGERS, KATHLEEN E  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING  
 ; TITLE OF INVENTION: ADHESION FORMATION  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ROBBINS, DALGARN, BERLINER & CARSON  
 ; STREET: 201 NORTH FIGUEROA STREET, FIFTH FLOOR  
 ; CITY: LOS ANGELES  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 90012-2628  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/178,482  
 ; FILING DATE: 06-JAN-1994  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/789,231  
 ; FILING DATE: 07-NOV-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: SPITALS, JOHN P  
 ; REGISTRATION NUMBER: 29,215  
 ; REFERENCE/DOCKET NUMBER: 1920-314  
 ; TELEPHONE: (213) 977-1001  
 ; TELEFAX: (213) 977-1003  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear

## US-08-178-482-12

Query Match 100.0%; Score 27; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 6

US-08-259-569-7  
 ; Sequence 7, Application US/08259569  
 ; Patent No. 5679320  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vogel, Tikva  
 ; APPLICANT: Levanon, Avigdor  
 ; APPLICANT: Werber, Moshe  
 ; APPLICANT: Guy, Rachel  
 ; APPLICANT: Panet, Amos

APPLICANT: Hartman, Jacob  
 APPLICANT: Shaked, Hadassa  
 TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND  
 TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME  
 NUMBER OF SEQUENCES: 39  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/259,569  
 FILING DATE:  
 CLASSIFICATION: 436  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 703,842  
 FILING DATE: 21-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-977-9550  
 TELEFAX: 212-664-0525  
 TELEX: 422523 COOP UI  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal

## US-08-259-569-7

Query Match 100.0%; Score 27; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 7

US-08-240-711-9  
 ; Sequence 9, Application US/08240711  
 ; Patent No. 5679777  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDERSON, DAVID C.  
 ; APPLICANT: MATHEWS, ANTHONY J.  
 ; TITLE OF INVENTION: HEMOGLOBINS AS DRUG DELIVERY AGENTS  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Browdy and Neimark  
 ; STREET: 419 Seventh Street, N.W., Suite 300  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20004  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

## US-08-240-711-9

Query Match 100.0%; Score 27; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,711  
FILING DATE: 12-JUL-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA: PCT/US92/59713  
FILING DATE: 06-NOV-1992  
PRIOR APPLICATION DATA: US 07/789,177  
FILING DATE: 08-NOV-1991  
PRIOR APPLICATION DATA: US 07/789,179  
FILING DATE: 08-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=5A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-240-711-9

Query Match 100.0%; Score 27; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
|||  
DB 1 GRGDS 5

RESULT 8  
US-08-173-116-1  
Sequence 1, Application US/08173116  
Patent No. 5695761  
GENERAL INFORMATION:  
APPLICANT: Denhardt, David T.  
APPLICANT: Hwang, Shiaw-min  
APPLICANT: Heck, Diane E.  
APPLICANT: Lopez, Cecilia A.  
APPLICANT: Laskin, Jeffrey D.  
APPLICANT: Laskin, Debra L.  
TITLE OF INVENTION: SUPPRESSION OF NITRIC OXIDE PRODUCTION  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/173,116  
FILING DATE: 23-DEC-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 960-1-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-173-116-1

Query Match 100.0%; Score 27; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
|||  
DB 1 GRGDS 5

RESULT 9  
US-08-457-753-9  
Sequence 9, Application US/08457753  
Patent No. 5759512  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTHONY J.  
TITLE OF INVENTION: HEMOGLOBINS AS DRUG DELIVERY AGENTS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Neimark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,753  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,711  
FILING DATE: 12-JUL-1994  
APPLICATION NUMBER: PCT/US92/09713  
FILING DATE: 06-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/789,177  
FILING DATE: 08-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/789,179  
FILING DATE: 08-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=5A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 9:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-457-753-9

Query Match      100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 10
US-08-482-106-18
; Sequence 18, Application US/08482106
; Patent No. 5767071
; GENERAL INFORMATION:
; APPLICANT: Palladino, Michael A.
; APPLICANT: Lee, Bruce A.
; APPLICANT: Huse, William D.
; APPLICANT: Varner, Judith A.
; TITLE OF INVENTION: Sevenmer Cyclic Peptide Inhibitors of Diseases
; TITLE OF INVENTION: Involving Alpha(v)Beta(3)
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,106
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, Tom M.
; REGISTRATION NUMBER: 26,314
; REFERENCE/DOCKET NUMBER: IXYS-0002/00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5000
; TELEFAX: 415-857-0663
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-106-18

Query Match      100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 11
US-08-482-107B-1
; Sequence 1, Application US/08482107B
; Patent No. 5780426
; GENERAL INFORMATION:
; APPLICANT: Palladino, Michael A.
; APPLICANT: Lee, Bruce A.
; APPLICANT: Huse, William D.
; APPLICANT: Varner, Judith A.
; TITLE OF INVENTION: Fivemer Cyclic Peptide Inhibitors of
; TITLE OF INVENTION: Diseases Involving Alpha(v)Beta(3)
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: 127 Peachtree Street, N.E., Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,107B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 09051.0001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404 688 0770
; TELEFAX: 404 688 9880
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-107B-1

Query Match      100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 12
US-08-462-720-16
; Sequence 16, Application US/08462720
; Patent No. 5849701
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D.
; APPLICANT: Kruttsch, Henry C.
; APPLICANT: Sipes, John M.
; APPLICANT: Guo, Neung-hua
; APPLICANT: Negre, Eric
; TITLE OF INVENTION: Peptide Inhibitors of Fibrinectin and
; TITLE OF INVENTION: Related Collagen-Binding Proteins
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/462,720
? FILING DATE: 05-JUN-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Dow, Karen B.
? REGISTRATION NUMBER: 29,684
? REFERENCE/DOCKET NUMBER: 015280-0231110US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 326-2400
? TELEFAX: (415) 326-2422
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-08-462-720-16

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
Db 1 GRGDS 5

RESULT 13
US-08-826-885-7
? Sequence 7, Application US/08326885
? Patent No. 5869616
? GENERAL INFORMATION:
? APPLICANT: Vogel, Tikva
? APPLICANT: Levanon, Avigdor
? APPLICANT: Werber, Moshe
? APPLICANT: Guy, Rachel
? APPLICANT: Panet, Amos
? APPLICANT: Hartman, Jacob
? APPLICANT: Shaked, Hadassa
? TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
? TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME
? NUMBER OF SEQUENCES: 39
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Cooper & Dunham
? STREET: 30 Rockefeller Plaza
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10112
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? FILING DATE:
? APPLICATION NUMBER: JS/08/826,885
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/703,842
? FILING DATE: 21-MAY-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: White, John P.
? REGISTRATION NUMBER: 28,678
? REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-977-9550
? TELEFAX: 212-664-0525
? TELEX: 422523 COOP UI
```

```
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal
? US-08-826-885-7

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
Db 1 GRGDS 5

RESULT 14
US-08-747-137-71
? Sequence 71, Application US/08747137
? Patent No. 5945033
? GENERAL INFORMATION:
? APPLICANT: Yen, Richard C.K.
? TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
? TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE
? NUMBER OF SEQUENCES: 184
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Townsend and Townsend and Crew LLP
? STREET: Two Embarcadero Center, 8th Floor
? CITY: San Francisco
? STATE: CA
? COUNTRY: USA
? ZIP: 94111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/747,137
? FILING DATE: 12-NOV-1996
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/212,546
? FILING DATE: 14-MAR-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/069,831
? FILING DATE: 01-JUN-1993
? APPLICATION DATA:
? APPLICATION NUMBER: US 07/959,560
? FILING DATE: 13-OCT-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/641,720
? FILING DATE: 15-JAN-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Apple, Randolph T.
? REGISTRATION NUMBER: 36,429
? REFERENCE/DOCKET NUMBER: 016197-000840US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-576-0200
? INFORMATION FOR SEQ ID NO: 71:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: not relevant
? US-08-747-137-71

Query Match 100.0%; Score 27; DB 2; Length 5;
```

Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
Db 1 GRGDS 5

RESULT 15  
US-08-747-137-181  
Sequence 181, Application US/08747137  
Patent No. 5945033  
GENERAL INFORMATION:  
APPLICANT: YEN, Richard C.K.  
TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR  
THERAPEUTIC AND DIAGNOSTIC USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew LLP  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,137  
FILING DATE: 12-NOV-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,546  
FILING DATE: 14-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/069,831  
FILING DATE: 01-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/959,560  
FILING DATE: 13-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/641,720  
FILING DATE: 15-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 016197-000840US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
INFORMATION FOR SEQ ID NO: 181:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
US-08-747-137-181

Query Match 100.0%; Score 27; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
Db 1 GRGDS 5

Search completed: April 16, 2004, 07:31:38  
Job time : 17.125 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw mode.

Run on: April 16, 2004, 06:55:57 ; Search time 54.625 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588B-4  
Perfect score: 27  
Sequence: 1 GRGDS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	27	100.0	5	2	AAR07442	Peptide w
2	27	100.0	5	2	AAR04610	Antiviral
3	27	100.0	5	2	AAR24516	Platelet
4	27	100.0	5	2	AAR27030	Peptide 1
5	27	100.0	5	2	AAR29051	Peptide 1
6	27	100.0	5	2	AAR36714	Adhesion
7	27	100.0	5	2	AAR34451	Fibronect
8	27	100.0	5	2	AAR44045	RGD pepti
9	27	100.0	5	2	AAR37131	RGD pepti
10	27	100.0	5	2	AAR17467	Tumour me
11	27	100.0	5	2	AAR35463	Propene-a
12	27	100.0	5	2	AAR5076	Fibronect
13	27	100.0	5	2	AAR48654	RGD conta
14	27	100.0	5	2	AAR62948	RGD contg
15	27	100.0	5	2	AAR70477	Cancer me
16	27	100.0	5	2	AAR25177	RGD-pepti
17	27	100.0	5	2	AAR65001	Synthetic
18	27	100.0	5	2	AAR51241	Peptide w
19	27	100.0	5	2	AAR80779	Fibronect
20	27	100.0	5	2	AAR37782	Fibronect
21	27	100.0	5	2	AAR43306	Epitope f
22	27	100.0	5	2	AAR99777	Cell atta
23	27	100.0	5	2	AAR03779	Human ost
24	27	100.0	5	2	AAR31132	Non-cross
25	27	100.0	5	2	AAR31022	Non-cross

ALIGNMENTS

RESULT 1  
AAR07442  
ID AAR07442 standard; peptide; 5 AA.  
XX AC AAR07442;  
XX AC AAR07442;  
DT 25-MAR-2003 (revised)  
DT 03-FEB-1991 (first entry)  
XX  
DB Peptide with anti-metastasis activity.  
XX Metastasis; cancer.  
XX Synthetic.  
OS Key Location/Qualifiers  
FH Misc-difference 2 /label= Side chain  
FT /note= "t-butyloxycarbonyl or sialic acid deriv. or H"  
FT Misc-difference 4 /label= Side chain  
FT /note= "Benzyloxycarbonyl or H"  
XX JP02233696-A.  
XX 17-SEP-1990-  
XX 06-MAR-1989; 89JP-00053496.  
XX 06-MAR-1989; 89JP-00053496.  
XX (NECT-) NECT CORP.  
XX WPI; 1990-325187/43.  
XX Penta peptide derivs. - and their prepn. for prevention of cancer metastasis.  
XX Claim 1; Page 1; 8pp; Japanese.  
XX Peptide is useful in prevention of cancer metastasis and has a high bioavailability. (Updated on 25-MAR-2003 to correct PA field.)  
SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 2
AAR04610
ID      AAR04610 standard; protein; 5 AA.
XX
AC      AAR04610;
XX
XX      25-MAR-2003 (revised)
DT      05-SEP-1990 (first entry)
XX
DE      Antiviral agent.
XX
KW      Antiviral; M2; poliovirus; polio; hepatitis.
XX
OS      Synthetic.
XX
XX      JP0207863-A.
XX      PN
XX      19-MAR-1990.
XX
XX      14-SEP-1988; 88JP-00228843.
XX      PF
XX      14-SEP-1988; 88JP-00228843.
XX      PR
XX      (NIHA ) NIPPON MINING CO.
XX      PA
XX      WPI; 1990-129060/17.
XX      DR
XX      Antiviral agent contg. tripeptide (unit) - of basic aminoacid, then
XX      PT alanine, glycine or sarcosine, and acidic aminoacid, effective against
XX      PT virus with protein-terminated DNA or RNA.
XX      XX
XX      Disclosure; Page ?; 4pp; Japanese.
XX      PS
XX      Peptide is effective against inhibiting propagation of DNA or RNA bonded,
XX      CC protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on
XX      CC 25-MAR-2003 to correct PA field.)
XX      CC
XX      Sequence 5 AA;

Query Match      100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 3
AAR24516
ID      AAR24516 standard; protein; 5 AA.
XX
AC      AAR24516;
XX
XX      02-DEC-1992 (first entry)
DT
DE      Platelet antagonist peptide 3.
XX
XX      Clinical effect; antagonist.
XX
XX      Synthetic.
XX
XX      JP04134096-A.
XX      PN
XX      07-MAY-1992.
XX      PD
XX      21-SEP-1990; 90JP-00253849.
XX      PF

Query Match      100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDS 5
Db      1 GRGDS 5

RESULT 4
AAR27030
ID      AAR27030 standard; peptide; 5 AA.
XX
XX      AAR27030;
XX
XX      20-MAY-1998 (first entry)
DT
XX      Peptide lipid contg. RGD.
XX
XX      Synthetic; cell migration; inhibitor; cell adhesion membrane;
XX      KW cell culture body.
XX
XX      Synthetic.
XX
XX      Key      Location/Qualifiers
XX      modified_site 1
XX      modified_site 5 /note= "acylated"
XX      modified_site 5 /note= "alkylated"
XX
XX      JP04221395-A.
XX
XX      11-AUG-1992.
XX      PD
XX      29-NOV-1990; 90JP-00333336.
XX      PF
XX      26-OCT-1990; 90JP-00289494.
XX      PR
XX      (FUJIF ) FUJIFILM PHOTO FILM CO LTD.
XX      PA
XX      WPI; 1992-313679/38.
XX      DR
XX      New synthetic peptide lipid(s) and salts - useful as cell migration
XX      PT inhibitors, cell adhesion membranes or cell culture bodies.
XX
XX      Disclosure; Page 3; 9pp; Japanese.
XX      PS
XX      The peptide sequence is an example of a highly generic sequence contg.
XX      CC the RGD motif. Compounds contg. these lipid peptides are useful as cell
XX      CC migration inhibitors in cell adhesion membranes or cell culture bodies.
XX      CC See also AA527027-33
XX      CC
XX      Sequence 5 AA;

```

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 1 GRGDS 5

RESULT 5  
 AAR29051  
 ID AAR29051 standard; peptide; 5 AA.  
 XX AC  
 XX AAR29051;  
 XX 20-MAY-1998 (first entry)  
 DT XX  
 DE XX  
 XX Peptide lipid contg. RGD.  
 KW Synthetic; cell migration; inhibitor; cell adhesion membrane;  
 KW cell culture body.  
 XX Synthetic.  
 OS XX  
 FH Key Location/Qualifiers  
 FT modified\_site 1 /note= "acylated"  
 FT modified\_site 5 /note= "alkylated"  
 FT  
 XX JP04221394-A.  
 XX 11-AUG-1992.  
 PD XX  
 XX 29-NOV-1990; 90JP-00333335.  
 XX 26-OCT-1990; 90JP-00289493.  
 PR XX  
 XX (FUJIFILM) FUJIFILM PHOTO FILM CO LTD.  
 PA WPI; 1992-313678/38.  
 DR XX  
 XX New synthetic peptide lipids or salts - useful as cell migration  
 FT inhibitors, cell adhesion membranes and cell culture bodies.  
 FT  
 XX Disclosure; Page 3; 9pp; Japanese.  
 ES  
 XX The peptide sequence is an example of a highly generic sequence contg.  
 CC the RGD motif. Compounds contg. these lipid peptides are useful as cell  
 CC migration inhibitors in cell adhesion membranes or cell culture bodies.  
 CC See also AAR29048-54  
 XX  
 XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 1 GRGDS 5

RESULT 6  
 AAR36714  
 ID AAR36714 standard; peptide; 5 AA.  
 XX AC  
 XX AAR36714;  
 XX 25-MAR-2003 (revised)  
 DT 26-AUG-1993 (first entry)  
 PD XX  
 XX

DE Adhesion formation prevention RGD-contg. peptide.  
 XX Tissue repair; peritoneum; surgery; post-surgically; inhibition;  
 KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;  
 KW CNS; use.  
 XX Synthetic.  
 OS  
 XX WO9308818-A1.  
 PN 13-MAY-1993.  
 PD XX  
 XX 06-NOV-1992; 92WO-US009494.  
 PE XX  
 XX 07-NOV-1991; 91US-00789231.  
 PR XX  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 PA Dizerega GS, Rodgers KE;  
 PI WPI; 1993-167381/20.  
 XX Prevention of adhesion formation, partic. post-surgically - comprises  
 PT administering a RGD-contg. peptide for a time sufficient to permit tissue  
 PT repair.  
 XX Example; Page 18; 22pp; English.  
 PS The sequence is that of an RGD-contg. peptide which is used in a method  
 XX for prevention of adhesion formation for a time sufficient to permit  
 CC tissue repair. The method is used for minimizing or preventing adhesion  
 CC formation, partic. in the peritoneum following surgery, but also for e.g.  
 CC cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In  
 CC addn., the peptide inhibits platelet aggregation and does not induce  
 CC inflammation or trauma at the site of administration. (Updated on 25-MAR-  
 CC 2003 to correct PN field.)  
 XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 1 GRGDS 5

RESULT 7  
 AAR34451  
 ID AAR34451 standard; peptide; 5 AA.  
 XX AC  
 XX AAR34451;  
 XX 25-MAR-2003 (revised)  
 DT 02-AUG-1993 (first entry)  
 DE Fibronectin-like binding motif.  
 XX Immunoglobulin D; oligosaccharide; N-glycosylation;  
 KW Igd antigen receptor complex; delta chain constant region.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Binding-site 2, 4  
 FT /note= "fibrinogen binding motif"  
 XX WO9307291-A1.  
 PN 15-APR-1993.  
 PD XX  
 XX 09-OCT-1992; 92WO-US008724.  
 PF

XX PR 11-OCT-1991; 91US-00773328.  
 XX PA (UUNY ) UNIV NEW YORK STATE.  
 XX PT Thorbecke GJ, Amin AR, Oppenheim JD;  
 XX DR WPI; 1993-134476/16.  
 XX  
 XX Use of immunoglobulin D-associated glycan cpds. - for inhibiting immune  
 XX response, treating immune-mediated disease and enhancing immune response.  
 XX  
 XX Disclosure; Page 53; 77pp; English.  
 XX  
 XX This pentapeptide sequence comprises the RGD motif typical of cell  
 XX adhesion molecules, including fibronectin. (Updated on 25-MAR-2003 to  
 XX correct PN field.)  
 XX SQ Sequence 5 AA;  
 Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 Db |||||  
 1 GRGDS 5  
 RESULT 8  
 AAR44045  
 ID AAR44045 standard; peptide; 5 AA.  
 XX AC AAR44045;  
 XX DT 02-JUN-1994 (first entry)  
 XX DE RGD peptide derivative #3.  
 XX  
 XX Drug; organ transplantation; rejection; immune disorder; systemic lupus.  
 XX  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 XX Modified-site 1  
 XX FT /note= "C15H31CO-Gly"  
 XX FT Modified-site 4  
 XX FT /note= "Ser-OH"  
 XX  
 XX JP05255105-A.  
 XX  
 XX 05-OCT-1993.  
 XX  
 XX PF 16-MAR-1992; 92JP-00058460.  
 XX  
 XX PR 16-MAR-1992; 92JP-00058460.  
 XX  
 XX (FUJF ) FUJI PHOTO FILM CO LTD.  
 XX  
 XX WPI; 1993-348360/44.  
 XX  
 XX Immuno-control drug for organ transplant rejection etc. - contains  
 XX peptide having arginine, glycine, aspartic acid sequence.  
 XX  
 XX Disclosure; Page 3; 11pp; Japanese.  
 XX  
 XX The sequences given in AAR44043-47 and AAR53144 represent examples of the  
 XX claimed RGD containing peptide of the invention. These peptides all  
 XX correspond to the generic formulae: HO2-(CH2)m-C(O)-([X]-Arg-Gly-Asp-  
 XX [Y])n-O-CH2CH(OR1)CH2OR2 or R3-([X]-Arg-Gly-Asp-[Y])n-Z [X] [Y] = amino  
 XX acid or peptide residues; m = 1-5; n = 1-5; R1, R2 = H or 8-24C acyl or  
 XX alkyl; R3 = 6-24C acyl; Z = hydroxyl or amino. These peptides form the  
 XX active part of drugs which are used for the control of organ

CC transplantation rejection or immune disorders such as systemic lupus  
 XX  
 XX SQ Sequence 5 AA;  
 Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 Db |||||  
 1 GRGDS 5  
 RESULT 9  
 AAR37131  
 ID AAR37131 standard; peptide; 5 AA.  
 XX AC AAR37131;  
 XX DT 19-APR-1994 (first entry)  
 XX DE RGD peptide deriv. #2.  
 XX  
 XX Cell adhesion core sequence; inhibitor; cancer; metastasis; ss.  
 XX  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 XX Modified-site 1  
 XX FT /label= Suc-Gly  
 XX  
 XX JP05222092-A.  
 XX  
 XX PD 31-AUG-1993.  
 XX  
 XX PF 16-JUL-1992; 92JP-00189446.  
 XX  
 XX PR 18-DEC-1991; 91JP-00335213.  
 XX  
 XX (FUJF ) FUJI PHOTO FILM CO LTD.  
 XX  
 XX WPI; 1993-309192/39.  
 XX  
 XX Peptide derivs. used as cancer metastasis inhibitors - contain the cell  
 XX adhesion core sequence Arg-Gly-Asp followed by Ser.  
 XX  
 XX Disclosure; Page 3; 10pp; Japanese.  
 XX  
 XX The sequences given in AAR37130-48 are peptide derivatives which  
 XX correspond to the generic formula; R1-X-Arg-Gly-Ser-R2; R1 = H or an  
 XX optionally substituted acyl, R2 = -OR3 or -NR4R5, R3, R4, R5 = H or 1-3C  
 XX alkyl, and X = a bond or an amino acid or peptide residue. These peptide  
 XX derivs. contain the cell adhesion core sequence Arg-Gly-Asp. They are  
 XX useful as inhibitors of cancer metastasis  
 XX  
 XX SQ Sequence 5 AA;  
 Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 Db |||||  
 1 GRGDS 5  
 RESULT 10  
 AAY17467  
 ID AAY17467 standard; peptide; 5 AA.  
 XX AC AAY17467;  
 XX  
 XX 30-JUL-1999 (first entry)

XX DE Tumour metastasis inhibitor peptide #2.  
 XX KW Tumour metastasis inhibitor; intercellular adhesion.  
 XX OS Synthetic.  
 XX PN JP05163300-A.  
 XX PD 29-JUN-1993.  
 XX PF 17-DEC-1991; 91JP-00333543.  
 XX PR 17-DEC-1991; 91JP-00333543.  
 XX PA (FUJIF) FUJIFILM PHOTO FILM CO LTD.  
 XX WI 1993-240027/30.  
 XX PT Biologically active peptide deriv. contg. arginine, glycine and aspartic acid units - used as tumour metastasis inhibitor having high intercellular adhesion.  
 XX PS Example 5; Page 11; 12pp; Japanese.  
 XX CC The present invention describes peptide derivatives of formula Q-W-CO-(X)-Arg-Gly-Asp-(Y)-R (I) and the pharmaceutically acceptable salts. In (I), W = 1-15C straight or substituted alkylene or arylene, optionally interrupted by -O-, -NH-, -S-, ester bond, amide bond, urethane bond or urea bond; X = Gly residue or is absent; Y = amino acid residue or peptide residue selected from Ser, Thr, Asp, Val and Ser-Pro, or is absent; R = -OH or NR1R2 R1 and R2 = H or 1-6C alkyl, or may together form a ring; and Q = acid sugar residue (excluding sialic acid residue) linked to W by a glycoside bond. Q is a sugar residue containing carboxyl, sulphonyl and/or phosphoryl group on the sugar chain; i.e. sugar residue selected from uronic acid, oligosaccharide comprising uronic acid, 6-O-carboxy-methyl-D-glucose, oligosaccharide comprising glucuronic acid and galactose, oligosaccharide comprising glucuronic acid and N-acetyl-glucosamine, oligosaccharide comprising glucuronic acid and galactosamine, oligosaccharide comprising glucuronic acid and W-acetyl galactosamine. (I) has great intercellular adhesion compared with a core sequence of fibronectin and sufficient biological activities but little toxicity. The present sequence represents an example of (I)

XX SQ Sequence 5 AA;  
 Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 |||||  
 DB 1 GRGDS 5

RESULT 11  
 AAR35463  
 ID AAR35463 standard; peptide; 5 AA.  
 AC AAR35463;  
 XX 26-AUG-1993 (first entry)  
 DE Propene-amide deriv. polymer metastasis inhibitor.  
 XX Low toxicity; higher cell adhesion ability; metastasis inhibition.  
 XX OS Synthetic.  
 XX Key Location/Qualifiers  
 XX Modified-site 1  
 FT /note= "CH2=CCH3CONHCH2-CH2-CO-Gly, CH2=CC2H5CONH-(CH2)4-

FT CO-Gly or CH2=CCH3CONHCH2-CH2-O-CH2-CO-Gly"  
 XX JP05097699-A.  
 XX 20-APR-1993.  
 XX PF 04-OCT-1991; 91JP-00258095.  
 XX PR 04-OCT-1991; 91JP-00258095.  
 XX PA (FUJIF) FUJIFILM PHOTO FILM CO LTD.  
 XX WI 1993-164370/20.  
 XX PT Low toxicity metastasis inhibitor - composed of propene-amide deriv. polymer or its pharmacologically acceptable salts.  
 XX PS Claim 1; Page 2; 12pp; Japanese.  
 XX CC The sequence is that of a polymer of propene amide deriv. which has a higher cell adhesion ability, compared with that of the core sequence of cell adhesive protein. It has various kinds of biological activities e.g. metastasis inhibition and has low toxicity

QY 1 GRGDS 5  
 |||||  
 DB 1 GRGDS 5

RESULT 12  
 AAR55076  
 ID AAR55076 standard; peptide; 5 AA.  
 AC AAR55076;  
 XX 25-MAR-2003 (revised)  
 DT 16-DEC-1994 (first entry)  
 DE Fibronectin gelatin binding domain inhibitory peptide.  
 XX fibronectin; collagen-binding proteins; inhibit cell adhesion;  
 XX regulate cell matrix interactions; tumorigenesis; metastasis;  
 XX wound repair; homostasis; thrombospondin.  
 XX OS Synthetic.  
 XX WO9411395-A1.  
 XX 26-MAY-1994.  
 XX PF 09-NOV-1993; 93WO-US011104.  
 XX PR 10-NOV-1992; 92US-00973235.  
 XX (USSH) US SEC DEPT HEALTH.  
 XX Roberts DE, Krutcz HC, Sipes JM, Guo N, Negre E;  
 WI 1994-183422/22.  
 PT Peptides which bind to fibronectin and collagen-binding proteins - are used to inhibit fibronectin dependent cell adhesion to collagen matrices.  
 XX Disclosure; Page 36; 51pp; English.  
 XX Thrombospondin is a multi-functional protein capable of interacting with numerous molecules, eg. fibronectin. Peptides have been designed that are

CC derived from the second type I repeat of human endothelial cell  
 CC thrombospondin. The peptides can be used to bind to fibronectin or other  
 CC related collagen-binding proteins to inhibit fibronectin-dependent cell  
 CC adhesion to collagen matrices. The peptides interact directly with the  
 CC gelatin-binding domain of fibronectin and inhibits the fibronectin  
 CC function. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-  
 CC MAR-2003 to correct PR field.)  
 XX  
 XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 |||||  
 Db 1 GRGDS 5

# RESULT 13

AAR48654  
 ID AAR48654 standard; peptide; 5 AA.

AC AAR48654;

XX 25-MAR-2003 (revised)  
 DT 21-SEP-1994 (first entry)  
 XX

DE RGD containing peptide #3 for attaching cells to a solid carrier.

KW RGD; solid substrate; solid carrier; attachment; culture;  
 KW anchorage dependant cells; receptor; cell surface; release.

XX Synthetic.

XX BP587205-A1.

XX 16-MAR-1994.

XX 19-JUL-1993; 93EP-00202102.  
 XX

PR 23-JUL-1992; 92EP-00202251.

PA (DUIN ) DUFAR INT RES BV.

XX Brands R, Snoek MC;

XX WPI; 1994-085139/11.

PT Solid carrier for culture of anchorage dependent cells - non-covalently  
 PT coated with a polypeptide contained specific sequence for easy attachment  
 PT and detachment of cells.

XX Example 1; Page 3; 7pp; English.

CC The sequences given in AAR48652-54 are RGD containing peptides which are  
 CC used to coat a solid substrate in the production of a solid carrier  
 CC suitable for the attachment of cells. The peptide molecules are directly  
 CC bound to the substrate by a non-covalent linkage. The solid carrier may  
 CC be used for culturing anchorage dependant cells. The epitope RGD is  
 CC recognised by receptors on the cell surface for cell attachment. The  
 CC cells are subsequently released from the carrier by simply lowering the  
 CC pH of the growth medium to below 7, optionally in the presence of  
 CC trypsin. The ready release of the cells from the solid carrier minimises  
 CC the damage to cell surface proteins. (Updated on 25-MAR-2003 to correct  
 CC PN field.)  
 XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 |||||  
 Db 1 GRGDS 5

# RESULT 14

AAR62948  
 ID AAR62948 standard; peptide; 5 AA.

AC AAR62948;

XX 15-AUG-1995 (first entry)  
 DT

DE RGD contg. peptide is a constituent of a cancer metastasis inhibitor.

KW RGD peptide; Arg-Gly-Asp; cancer metastasis inhibitor;  
 KW high cell adhesion; fibronectin.

XX Synthetic.

XX JF06306096-A.

XX 01-NOV-1994.

XX 31-JAN-1994; 94JP-00009893.

XX 26-FEB-1993; 93JP-00038677.

XX (FUJF ) FUJI PHOTO FILM CO LTD.

XX WPI; 1995-019276/03.

PT Peptide deriv. comprising several copies of the sequence Arg-Gly-Asp -  
 PT covalently bonded to organic molecule, useful as cancer metastasis  
 PT inhibitor.

XX Disclosure; Page 3; 13pp; Japanese.

CC AAR62945-50 are RGD contg. peptides that form constituents of an organic  
 CC mol. with a defined copy number of one of the peptides. The organic mol.  
 CC contg. the RGD peptide is a cancer metastasis inhibitor. The deriv. shows  
 CC substantially no toxicity and has a higher cell adhesion than the core  
 CC sequence of fibronectin  
 XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 |||||  
 Db 1 GRGDS 5

# RESULT 15

AAR70477  
 ID AAR70477 standard; peptide; 5 AA.

AC AAR70477;

XX 20-DEC-1995 (first entry)  
 DT

DE Cancer metastasis inhibitory RGD peptide derivative #5.

KW Cancer metastasis; adhesive peptide; core sequence; dextran; cancer;  
 KW water soluble polysaccharide; metastasis; wound; immunogenicity.

XX Synthetic.

XX JF07089999-A.

XX 04-APR-1995.

XX 17-SEP-1993; 93JP-00254779.  
 XX 17-SEP-1993; 93JP-00254779.  
 XX (JAPG ) NIPPON ZEON KK.  
 XX WPI; 1995-167254/22.  
 DR  
 XX  
 XX Cancer metastasis inhibitive peptide derivs. - useful for inhibition of  
 PT cancer metastasis, healing of wounds and regulation of immunogenicity.  
 XX  
 XX Disclosure; Page 2; 6pp; Japanese.  
 XX  
 CC The peptides AAR70472-90 and AAR82907-24 are peptide derivatives which  
 CC inhibit cancer metastasis. They are composed of an adhesive peptide with  
 CC a core sequence selected from: RGD (AAR70472-85), YIGSR (AAR70486-90) or  
 CC other sequence (AAR82907-24), linked to a water soluble polysaccharide,  
 CC preferably a water soluble dextran, at the C-terminus. The peptides are  
 CC useful in inhibiting cancer metastasis, healing wounds and the regulation  
 CC of immunogenicity  
 XX  
 SQ Sequence 5 AA;  
 Query Match 100.0%; Score 27; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 Db 1 GRGDS 5

Search completed: April 16, 2004, 07:20:50  
 Job time : 56.625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: April 16, 2004, 06:55:57 ; Search time 65.55 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588B-7  
Perfect score: 34  
Sequence: 1 GRGDSF 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	6	2	AAR04612
2	34	100.0	6	2	AAR06455
3	34	100.0	6	2	AAR29063
4	34	100.0	6	2	AAR29068
5	34	100.0	6	2	AAR24212
6	34	100.0	6	2	AAR29158
7	34	100.0	6	2	AAR29070
8	34	100.0	6	2	AAR29056
9	34	100.0	6	2	AAR29061
10	34	100.0	6	2	AAR26814
11	34	100.0	6	2	AAR26807
12	34	100.0	6	2	AAR27033
13	34	100.0	6	2	AAR25425
14	34	100.0	6	2	AAR24241
15	34	100.0	6	2	AAR22969
16	34	100.0	6	2	AAR29054
17	34	100.0	6	2	AAR36712
18	34	100.0	6	2	AAR36708
19	34	100.0	6	2	AAR36710
20	34	100.0	6	2	AAR32387
21	34	100.0	6	2	AAR37834
22	34	100.0	6	2	AAR32421
23	34	100.0	6	2	AAR44046
24	34	100.0	6	2	AAR35465
25	34	100.0	6	2	AAR37029

26	34	100.0	6	2	AAR57890
27	34	100.0	6	2	AAR49801
28	34	100.0	6	2	AAR70654
29	34	100.0	6	2	AAR79077
30	34	100.0	6	2	AAR71456
31	34	100.0	6	2	AAR70478
32	34	100.0	6	2	AAR80966
33	34	100.0	6	2	AAR99889
34	34	100.0	6	2	AAR94570
35	34	100.0	6	2	AAW01130
36	34	100.0	6	2	AAW07431
37	34	100.0	6	2	AAR92740
38	34	100.0	6	2	AAW03680
39	34	100.0	6	2	AAW03484
40	34	100.0	6	2	AAR88570
41	34	100.0	6	2	AAW25186
42	34	100.0	6	2	AAW25182
43	34	100.0	6	2	AAW15598
44	34	100.0	6	2	AAW79658
45	34	100.0	6	2	AAW34089

ALIGNMENTS

RESULT 1  
AAR04612  
ID AAR04612 standard; protein; 6 AA.  
AC AAR04612;  
XX  
XX  
DT 25-MAR-2003 (revised)  
DT 05-SEP-1990 (first entry)  
XX  
XX  
DE Antiviral agent.  
XX  
XX  
XX Antiviral; M2; poliovirus; polio; hepatitis.  
XX  
OS Synthetic.  
XX  
XX JPC2078631-A.  
XX  
PD 19-MAR-1990.  
XX  
XX  
PF 14-SEP-1988; 88JP-00228843.  
XX  
PR 14-SEP-1988; 88JP-00228843.  
XX  
XX (NIHA ) NIPPON MINING CO.  
XX  
XX WPI; 1990-129060/17.  
XX  
XX Antiviral agent contg. tripeptide (unit) - of basic aminoacid, then alanine, glycine or sarcosine, and acidic aminoacid, effective against virus with protein-terminated DNA or RNA.  
XX  
XX Disclosure; Page 7; 4pp; Japanese.  
XX  
XX Peptide is effective against inhibiting propagation of DNA or RNA bonded, protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSF 6  
DB 1 GRGDSF 6



```

RESULT 2
AAR06455
ID AAR06455 standard; peptide; 6 AA.
AC AAR06455;
XX
XX 25-MAR-2003 (revised)
DT 04-JAN-1991 (first entry);
XX
XX Fibronectin derived RGD-contg. peptide.
XX
XX Tenascin; receptor; tumour.
XX
XX Synthetic.
XX
XX WO9008781-A.
XX
XX 09-AUG-1990.
XX
XX 24-JAN-1989; 89US-00302755.
XX
XX 24-JAN-1989; 89US-00302755.
XX
XX (LJOL-) LA JOLLA CANCER RES FOUND.
XX (LJOL-) LA JOLLA CANCER RES FOUND.
XX
XX Ruoslahti E, Bourdon MA;
XX
XX WPI; 1990-260895/34.
XX
XX Tenascin-induced cell attachment - interacts with tenascin in arginine-
XX glycine-asparagine dependent manner.
XX
XX Example II; Page 9; 22pp; English.
XX
XX The peptide was used in cell attachment assays to show inhibition of
XX attachment to tenascin by Arg-Gly Asp contg. peptides. It completely
XX inhibited attachment at a concn. of 300ug/ml, 30- and 50-fold lower than
XX needed for comparable inhibition of cell attachment to vitronectin and
XX fibronectin resp. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 6 AA;
SQ
Query Match 100.0%; Score 34; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRGDSP 6
Db 1 GRGDSP 6
XX
RESULT 3
AAR29063
ID AAR29063 standard; peptide; 6 AA.
AC AAR29063;
XX
XX 20-MAY-1998 (first entry);
XX
XX Peptide contg. RGD motif as a side chain to a water sol. polymer.
XX
XX Adhesive peptide; cell adhesion; inhibitor; platelet aggregation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX modified_site 1
XX modified_site 6 /note= "alkylated/arylated"
XX modified_site 6 /note= "alkylated/arylated"
XX
XX JP04221396-A.
XX
XX 11-AUG-1992.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX (FUJF) FUJI PHOTO FILM CO LTD.
XX
XX WPI; 1992-313680/38.
XX
XX Water-soluble vinyl polymer deriv. - for animal cell adhesion inhibitor
XX or platelet aggregation-adhesion inhibitor.
XX
XX Example; Page 13; 14pp; Japanese.
XX
XX The peptide sequence contains the Arg-Gly-Asp motif of cell adhering
XX proteins. It comprises the essential unit of a water-sol. vinyl polymer

```

```

XX
XX 11-AUG-1992.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX (FUJF) FUJI PHOTO FILM CO LTD.
XX
XX WPI; 1992-313680/38.
XX
XX Water-soluble vinyl polymer deriv. - for animal cell adhesion inhibitor
XX or platelet aggregation-adhesion inhibitor.
XX
XX Example; Page 12; 14pp; Japanese.
XX
XX The peptide sequence contains the Arg-Gly-Asp motif of cell adhering
XX proteins. It comprises the essential unit of a water-sol. vinyl polymer
XX with a pref. mol. wt. of 3000-100,000 D. The polymer shows various
XX biological activities, e.g. immunological coordination, wound healing
XX action and platelet aggregation inhibiting action etc. See also AAR29062-
XX 8
XX
XX Sequence 6 AA;
SQ
Query Match 100.0%; Score 34; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRGDSP 6
Db 1 GRGDSP 6
XX
RESULT 4
AAR29068
ID AAR29068 standard; peptide; 6 AA.
AC AAR29068;
XX
XX 20-MAY-1998 (first entry)
XX
XX Peptide contg. RGD motif as a side chain to a water sol. polymer.
XX
XX Adhesive peptide; cell adhesion; inhibitor; platelet aggregation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX modified_site 1
XX modified_site 6 /note= "alkylated/arylated"
XX modified_site 6 /note= "alkylated/arylated"
XX
XX JP04221396-A.
XX
XX 11-AUG-1992.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX 20-DEC-1990; 90JP-00404347.
XX
XX (FUJF) FUJI PHOTO FILM CO LTD.
XX
XX WPI; 1992-313680/38.
XX
XX Water-soluble vinyl polymer deriv. - for animal cell adhesion inhibitor
XX or platelet aggregation-adhesion inhibitor.
XX
XX Example; Page 13; 14pp; Japanese.
XX
XX The peptide sequence contains the Arg-Gly-Asp motif of cell adhering
XX proteins. It comprises the essential unit of a water-sol. vinyl polymer

```

CC with a pref. mol. wt. of 3000-100,000 D. The polymer shows various  
 CC biological activities, e.g. immunological coordination, wound healing  
 CC action and platelet aggregation inhibiting action etc. See also AAR29062-  
 CC 7

SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 5  
 AAR24212  
 ID AAR24212 standard; protein; 6 AA.

XX AAR24212;  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 18-NOV-1992 (first entry)

Fragment of tenascin-related peptide.

XX Tenascin; related peptide; cell attachment; antibody; angiogenesis;  
 KW tumour metastasis; solid matrix; prosthetic device; vascular graft;  
 KW percutaneous device.

XX Synthetic.

PN W09207872-K1.

PD 14-MAY-1992.

PF 29-OCT-1991; 91WO-US008018.

PR 29-OCT-1990; 90US-00605920.

PR 30-OCT-1990; 90US-00605667.

XX (CALB-) CALIFORNIA INST BIOLOGICAL RES.

XX Bourdon MA;

WPI; 1992-183625/22.

XX New tenascin-related peptides - modulate cell attachment to tenascin,  
 PT useful in inhibition of tumour metastasis and angiogenesis.

XX Disclosure; Page 8; 60pp; English.

XX The peptide may form an N- or C-terminal fragment of the generic peptide  
 CC of AAR4192, which is a tenascin-related peptide. This peptide mimics the  
 CC ability of tenascin to promote cell attachment. The peptide and  
 CC antibodies raised to it can be used to modulate cell attachment to  
 CC tenascin, esp. to inhibit tumour metastasis and angiogenesis. The peptide  
 CC is pref. attached to a solid matrix, eg collagen, nitrocellulose,  
 CC polyester, glass, synthetic resin, long-chain polysaccharide or synthetic  
 CC resin fibre. It is esp. operatively linked to a solid matrix forming a  
 CC prosthetic device, percutaneous device, vascular graft, etc. For topical  
 CC admin. it is formulated into a lotion, saline, gel, colloid, powder etc.  
 CC (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003  
 CC to correct PN field.)

XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 6

AAR29158  
 ID AAR29158 standard; peptide; 6 AA.

XX AAR29158;

XX 15-APR-1993 (first entry)

DE PEG-contg. peptide deriv #1 comprising RGD-motif.

XX Cell adhesion; Polyethylene glycol; fibronectin.

XX Synthetic.

XX Key Modified-site Location/Qualifiers  
 FT 1

FT /note= "acylated by ROCCH2 (OCH2CH2)NOCH2CO- where R =  
 GRGDSP and n = 1-150"

FT Modified-site 6

FT /note= "opt. amidated and if so, R is also amidated"

XX JP04305597-A.

XX 28-OCT-1992.

PF 02-APR-1991; 91JP-00068669.

PR 02-APR-1991; 91JP-00068669.

XX (FUJIF) FUJIFILM CO LTD.

XX WPI; 1992-410149/50.

XX New peptide-contg. polyethylene glycol derivs. - used to inhibit cancer  
 PT metastasis or platelet aggregation and as lymphocyte activators.

XX Example 1; Page 3; 6pp; Japanese.

XX The polyethylene glycol-contg. peptide derivs. contain the Arg-Gly-Asp  
 CC (RGD) motif found in fibronectin. The derivs can be used as platelet  
 CC aggregation inhibitors, lymphocyte activators and cancer metastasis  
 CC inhibitors. See also AAR29159 and AAR33149

XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 7

AAR29070

ID AAR29070 standard; peptide; 6 AA.

XX AAR29070;

XX 20-MAY-1998 (first entry)

DE Gelatin deriv. peptide contg. RGD motif.

XX Adhesive peptide; cell adhesion; inhibitor; platelet aggregation.

XX Synthetic.

PE	Key	Location/Qualifiers	
FT	modified_site	1	
FT	modified_site	6	/note= "alkylated/arylated"
FT	modified_site	6	/note= "alkylated/arylated"
XX			
XX	JF04221400-A.		
XX			
XX	11-AUG-1992.		
PD			
XX			
PP	30-NOV-1990;	90JP-00333719.	
XX			
PR	26-OCT-1990;	90JP-00289492.	
XX			
XX	(FUUF ) FUJI PHOTO FILM CO LTD.		
XX			
XX	WPI; 1992-313683/38.		
XX			
PT	Gelatin deriv. with adhesive peptide side chain - used for animal cell		
PT	adhesion inhibitor and platelet aggregation-adhesion inhibitor.		
XX			
XX	Example; Page 12; 10pp; Japanese.		
XX			
CC	The gelatin deriv. contains the Arg-Gly-Asp motif of cell adhering		
CC	proteins. It comprises the essential unit of a water-sol. vinyl polymer		
CC	with a pref. mol. wt. of 3000-100,000 D. The polymer shows various		
CC	biological activities, e.g. immunological coordination, wound healing		
CC	action and platelet aggregation inhibiting action etc. See also AAR29069-		
CC	75		
XX			
XX	Sequence 6 AA;		
QQ			
DD			
	Query Match	100.0%; Score 34; DB 2; Length 6;	
	Best Local Similarity	100.0%; Pred. No. 1.4e+06;	
	Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GRGDSP 6		
DD	1 GRGDSP 6		
RESULT 8			
AAR29056			
ID	AAR29056 standard; peptide; 6 AA.		
XX			
XX	AAR29056;		
XX			
DT	20-MAY-1998 (first entry)		
XX			
DE	Peptide contg. RGD motif as a side chain to a water sol. polymer.		
XX			
KW	Adhesive peptide; cell adhesion; inhibitor; platelet aggregation.		
XX			
CS	Synthetic.		
XX			
FF	Key	Location/Qualifiers	
FT	modified_site	1	
FT	modified_site	6	/note= "alkylated/arylated"
FT	modified_site	6	/note= "alkylated/arylated"
FT			
FT			
XX			
FN	JP04221397-A.		
XX			
PD	11-AUG-1992.		
XX			
XX	20-DEC-1990;	90JP-00403484.	
XX			
PR	20-DEC-1990;	90JP-00404484.	
XX			
XX	(FUUF ) FUJI PHOTO FILM CO LTD.		
XX			
XX	WPI; 1992-313681/38.		
XX			

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 10  
 AAR26814  
 ID AAR26814 standard; peptide; 6 AA.  
 XX  
 AC AAR26814;  
 XX  
 DT 20-MAY-1998 (first entry)  
 XX  
 DE Polyethylene glycol derivative #5.  
 XX  
 KW triazine ring; methoxy-polyethyleneoxy group; fibronectin; vitronectin;  
 XX platelet adhesion; metastasis; neuropathy.  
 OS Synthetic.  
 XX  
 PN JP04217693-A.  
 XX  
 PD 07-AUG-1992.  
 XX  
 PF 30-NOV-1990; 90JP-00333717.  
 XX  
 PR 23-OCT-1990; 90JP-00285172.  
 XX  
 PA (FUJF ) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1992-312284/38.  
 XX  
 PT Polyethylene glycol derivs. contg. peptide(s) - inhibit cellular adhesion  
 PT for fibronectin or vitronectin and are used to inhibit agglutination or  
 PT adhesion of platelets.  
 XX  
 PS Disclosure; Page 4; 9pp; Japanese.  
 XX  
 CC The sequences given in AAR26810-14 are examples of a peptide chain which  
 CC is attached once or twice to a triazine ring which is also substituted  
 CC twice or once, respectively, with a methoxy-polyethyleneoxy group. These  
 CC peptides can be used to inhibit cellular adhesion to fibronectin or  
 CC vitronectin and they are useful as inhibitors for agglutination or  
 CC adhesion of platelets. They can also be useful as inhibitors for  
 CC metastasis of cancers, inhibitors of agglutination of platelets caused by  
 CC tumour cells in the blood capillaries, and drugs acting on neuropathy  
 XX  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 11  
 AAR26807  
 ID AAR26807 standard; peptide; 6 AA.  
 XX  
 AC AAR26807;  
 XX  
 DT 12-FEB-1993 (first entry)  
 XX  
 DE Proten-amido peptide copolymer.  
 XX  
 KW Tumour metastasis inhibitor; platelet aggregation; animal cell; adhesion;  
 XX wound healing; cell culture media.

OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "CH2-CCH3-CO-Gly"  
 XX  
 PN JP04213311-A.  
 XX  
 PD 04-AUG-1992.  
 XX  
 PF 29-MAR-1991; 91JP-00066159.  
 XX  
 PR 27-NOV-1990; 90JP-00324610.  
 XX  
 PA (FUJF ) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1992-305482/37.  
 XX  
 PT New copolymers of propen-amido peptide(s) - are tumour metastasis,  
 PT platelet aggregation and animal cell adhesion inhibitors also useful as  
 PT wound healing agents and cell culture media.  
 XX  
 PS Example; Page 6; 14pp; Japanese.  
 XX  
 CC The sequence is that of a propen-amido peptide copolymer, it and its  
 CC salts are water soluble and is useful as a tumour metastasis inhibitor, a  
 CC platelet aggregation inhibitor, an animal cell adhesion inhibitor, a  
 CC wound healing agent and cell culture media. See also AAR26803-R26808  
 XX  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 12  
 AAR27033  
 ID AAR27033 standard; peptide; 6 AA.  
 XX  
 AC AAR27033;  
 XX  
 DT 20-MAY-1998 (first entry)  
 XX  
 DE Peptide lipid contg. RGD.  
 XX  
 KW Synthetic; cell migration; inhibitor; cell adhesion membrane;  
 KW cell culture body.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT modified\_site 1  
 FT /note= "acylated"  
 FT modified\_site 6  
 FT /note= "alkylated"  
 XX  
 PN JP04221395-A.  
 XX  
 PD 11-AUG-1992.  
 XX  
 PF 29-NOV-1990; 90JP-00333336.  
 XX  
 PR 26-OCT-1990; 90JP-00289494.  
 XX  
 PA (FUJF ) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1992-313679/38.  
 XX

PT New synthetic peptide lipid(s) and salts - useful as cell migration  
 XX inhibitors, cell adhesion membranes or cell culture bodies.  
 PS Disclosure; Page 4; 9pp; Japanese.

CC The peptide sequence is an example of a highly generic sequence contg.  
 CC the RGD motif. Compounds contg. these lipid peptides are useful as cell  
 CC migration inhibitors in cell adhesion membranes or cell culture bodies.  
 CC See also AAR27027-32

XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

RESULT 13

AAR25425  
 ID AAR25425 standard; peptide; 6 AA.

XX AC AAR25425;  
 XX 14-MAY-2003 (revised)  
 DT 06-JAN-1993 (first entry)

XX Peptide deriv. contg. RGD motif.

XX Phosphodiester; liposomes; micelles; fibronectin; cell adhesion;  
 KW targeting; tumour metastasis; agglutination; platelets; lymphocytes.  
 XX Synthetic.

XX JF04164095-A.

XX 09-JUN-1992.

XX 26-OCT-1990; 90JP-00289490.

XX 26-OCT-1990; 90JP-00289490.

XX (FUJIF) FUJI PHOTO FILM CO LTD.

XX WFI; 1992-239950/29.

XX New peptide contg. arginine-glycine-aspartic acid sequence - useful in  
 PT prepn. of liposome or micelles used to suppress tumour metastasis, since  
 PT sequence is activation site of fibronectin.

XX Example 1; Page 4; 9pp; Japanese.

XX The peptide is part of a phosphodiester bond which also comprises a  
 CC hydrophobic organic gp. e.g. an isoprenoid or glycerolipid. The new  
 CC derivs. of the peptide contg. the Arg-Gly-Asp sequence are useful for the  
 CC prepn. of liposomes or micelles contg. the RGD sequence. The RGD sequence  
 CC is an activation site of fibronectin which is a cell adhesion mol. The  
 CC liposomes are useful for the suppression of tumour metastasis,  
 CC agglutination of platelets, and activation of lymphocytes. They are  
 CC useful for targeting anti-tumour drugs onto tumours. See also AAR25426.  
 CC (Updated on 14-MAY-2003 to correct PS field.)

XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 DB 1 GRGDSP 6

DB 1 GRGDSP 6

RESULT 14

AAR24241  
 ID AAR24241 standard; protein; 6 AA.

XX AAR24241;

XX 25-MAR-2003 (revised)

DT 01-DEC-1992 (first entry)

XX Activation independent ligand.

XX Platelet aggregation disorder; ligand binding; occupancy; competent.

XX Synthetic.

XX WO9208982-A1.

XX 29-MAY-1992.

XX 15-NOV-1991; 91WO-US008579.

XX 15-NOV-1990; 90US-00614723.

XX (SCRI) SCRIPPS RES INST.

XX Ginsberg MH;

XX WFI; 1992-200317/24.

XX Rapid characterisation of platelet aggregation disorders - by detecting  
 PT levels of activation and ligand-occupancy competent platelets, pref. by  
 PT flow cytometry after reaction with labelled antibodies.

XX Claim 9; Page 70; 76pp; English.

XX The peptide is an activation independent ligand which is used as part of  
 CC a method for characterising a platelet aggregation defect in a patient  
 CC where the defect is an activation, ligand binding, or post occupancy  
 CC defect. It forms a ligand-induced binding site on normal platelets which  
 CC can be used to indicate the presence of ligand occupancy competent  
 CC platelets in a sample. The method provides rapid characterisation of  
 CC defects (less than 30 min. using platelet-rich plasma or whole blood).  
 CC gives improved definition of the defects and requires only 0.5 ml of  
 CC sample, conventional aggregation methods require about 20-30 ml. See also  
 CC AAR24239-R24242. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6

DB 1 GRGDSP 6

RESULT 15

AAR22969  
 ID AAR22969 standard; protein; 6 AA.

XX AAR22969;

XX 25-MAR-2003 (revised)

DT 26-OCT-1992 (first entry)

XX Cell adhesive peptide #3 on new CM chitin derivative side chains.

XX CM-chitin; cell adhesion; coadhesion; wound healing; RGD peptide;  
 KW immunoregulating agents; platelet coagulation; platelet adhesion.

XX Synthetic.  
 XX EP482649-A.  
 XX 29-APR-1992.  
 XX 24-OCT-1991; 91EP-00118179.  
 XX 26-OCT-1990; 90JP-00289491.  
 XX 30-NOV-1990; 90JP-00333718.  
 XX 29-MAR-1991; 91JP-00066156.  
 XX (FUJIF) FUJI PHOTO FILM CO LTD.  
 XX Kojima M, Komazawa H;  
 XX WPI; 1992-142753/18.  
 XX New CM chitin derivs. contg. adhesive peptides - have cell-adhesive  
 PT protein (antagonistic activity used for immuno-regulation and inhibiting  
 PT blood-platelet coagulation.  
 XX Example; Page 5; 24pp; English.  
 XX This peptide is an example of a claimed generic sequence, and is present  
 CC on the sidechain of new CM-chitin derivatives. The peptide may be  
 CC prepared by liquid or solid phase peptide synthesis and then coupled to  
 CC CM-chitin or carboxylated CM-chitin by amide bond forming methods using  
 CC eg CNBr, acid azides, or water soluble carbodiimides. The sequence  
 CC contains the core sequence RGD of a cell adhesive protein and the CM-  
 CC chitin derivatives contg. it will adhere to cells through the core  
 CC sequence according to a mechanism similar to that for the cell adhesive  
 CC protein. The derivatives act as antagonists of the cell adhesive protein  
 CC and can be used as animal cell adhesion-inhibiting agents, wound healing  
 CC agents, immunoregulating agents or platelet coagulation/adhesion-  
 CC inhibiting agents. See also AAR22967-70. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX Sequence 6 AA;  
 SQ

Query Match 100.0%; Score 34; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 Db 1 GRGDSP 6

Search completed: April 16, 2004, 07:20:53  
 Job time : 66.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 19.35 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-7  
Perfect score: 34  
Sequence: 1 GRGDSF 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 339414

Minimum DB seq length: 3  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/protdata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/protdata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/protdata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/protdata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/protdata/2/iaa/6C COMB.pep.\*  
6: /cgn2\_6/protdata/2/iaa/6D COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	6	1 US-07-720-222-46	Sequence 46, Appl
2	34	100.0	6	1 US-07-780-790A-6	Sequence 6, Appl
3	34	100.0	6	1 US-07-932-200-11	Sequence 11, Appl
4	34	100.0	6	1 US-08-234-979-3	Sequence 3, Appl
5	34	100.0	6	1 US-08-251-927-13	Sequence 13, Appl
6	34	100.0	6	1 US-08-214-770-13	Sequence 13, Appl
7	34	100.0	6	1 US-08-212-186A-2	Sequence 2, Appl
8	34	100.0	6	1 US-08-169-743-4	Sequence 4, Appl
9	34	100.0	6	1 US-08-262-315-1	Sequence 1, Appl
10	34	100.0	6	1 US-08-425-238-7	Sequence 7, Appl
11	34	100.0	6	1 US-08-178-482-6	Sequence 6, Appl
12	34	100.0	6	1 US-08-264-759-3	Sequence 3, Appl
13	34	100.0	6	1 US-08-445-745-49	Sequence 49, Appl
14	34	100.0	6	1 US-08-608-697-1	Sequence 1, Appl
15	34	100.0	6	1 US-08-596-116A-69	Sequence 69, Appl
16	34	100.0	6	1 US-08-104-335-3	Sequence 3, Appl
17	34	100.0	6	1 US-08-421-702A-4	Sequence 4, Appl
18	34	100.0	6	1 US-08-482-106-16	Sequence 16, Appl
19	34	100.0	6	1 US-08-303-052A-4	Sequence 4, Appl
20	34	100.0	6	1 US-08-421-696A-4	Sequence 4, Appl
21	34	100.0	6	1 US-07-803-623B-12	Sequence 12, Appl
22	34	100.0	6	1 US-08-482-107B-3	Sequence 3, Appl
23	34	100.0	6	1 US-08-421-697A-4	Sequence 4, Appl
24	34	100.0	6	1 US-08-421-698A-4	Sequence 4, Appl
25	34	100.0	6	2 US-08-387-749-11	Sequence 11, Appl
26	34	100.0	6	2 US-08-520-535-15	Sequence 15, Appl
27	34	100.0	6	2 US-08-280-646-2	Sequence 2, Appl

28	34	100.0	6	2 US-08-445-133-2	Sequence 2, Appl
29	34	100.0	6	2 US-08-432-638-2	Sequence 2, Appl
30	34	100.0	6	2 US-08-635-572-1	Sequence 1, Appl
31	34	100.0	6	2 US-08-806-084-12	Sequence 12, Appl
32	34	100.0	6	2 US-08-421-635A-4	Sequence 4, Appl
33	34	100.0	6	2 US-08-473-025-3	Sequence 3, Appl
34	34	100.0	6	2 US-08-447-810-2	Sequence 2, Appl
35	34	100.0	6	2 US-08-585-281-7	Sequence 7, Appl
36	34	100.0	6	2 US-08-893-853-5	Sequence 5, Appl
37	34	100.0	6	2 US-08-625-695A-2	Sequence 2, Appl
38	34	100.0	6	2 US-08-717-169-16	Sequence 16, Appl
39	34	100.0	6	2 US-08-279-773-2	Sequence 2, Appl
40	34	100.0	6	2 US-08-723-789-1	Sequence 1, Appl
41	34	100.0	6	2 US-08-747-137-72	Sequence 72, Appl
42	34	100.0	6	2 US-08-747-137-175	Sequence 175, Appl
43	34	100.0	6	2 US-09-079-432-15	Sequence 15, Appl
44	34	100.0	6	2 US-08-286-861-21	Sequence 21, Appl
45	34	100.0	6	3 US-08-915-189-83	Sequence 83, Appl

## ALIGNMENTS

RESULT 1  
US-07-720-222-46  
; Sequence 46, Application US/07720222  
; Patent NO. 5190873  
; GENERAL INFORMATION:  
; APPLICANT: Lernhardt, Waldemar  
; APPLICANT: Bourdon, Mario  
; APPLICANT: Youderian, Phil  
; TITLE OF INVENTION: HYBRID PROTEINS CONTAINING BINDING SITES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bingham & Fitting  
; STREET: 11230 Sorrento Valley Road, Suite 200  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07720,222  
; FILING DATE: 19910621  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bingham, Douglas A.  
; REGISTRATION NUMBER: 32,457  
; REFERENCE/DOCKET NUMBER: CIB0012P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-587-3533  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-07-720-222-46

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6  
|||||

Query Match 100.0%; Score 34; DB 1; Length 6;



Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 5  
US-08-251-027-13  
; Sequence 13, Application US/08251027  
; Patent No. 5519005  
; GENERAL INFORMATION:  
; APPLICANT: Lider, Ofer  
; APPLICANT: Greenspoon, No. 5519005m  
; APPLICANT: Hershkoviz, Rami  
; APPLICANT: Alon, Ronen  
; TITLE OF INVENTION: A METHOD OF INHIBITION OF CELLULAR AND  
; TITLE OF INVENTION: MOLECULAR LEVEL BIOLOGICAL INTERACTIONS UTILIZING  
; TITLE OF INVENTION: NON-PEPTIDIC SURROGATES OF THE ARG-GLY-ASP SEQUENCE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kohn & Associates  
; STREET: 30500 No. 5519005thwestern Highway, Suite 410  
; CITY: Farmington Hills  
; STATE: Michigan  
; COUNTRY: U.S.  
; ZIP: 48334  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/251,027  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohn, Kenneth I.  
; REGISTRATION NUMBER: 30,955  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (810) 539-5050  
; TELEFAX: (810) 539-5055  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-251-027-13

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 6  
US-08-2-4-770-13  
; Sequence 13, Application US/08214770  
; Patent No. 5523209  
; GENERAL INFORMATION:  
; APPLICANT: Ginsberg, Mark H.  
; APPLICANT: O'Toole, Tim  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: INHIBITORS OF INTEGRIN  
; TITLE OF INVENTION: ACTIVATION  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 7  
US-08-212-186A-2  
; Sequence 2, Application US/08212186A  
; Patent No. 5536814  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki I.  
; APPLICANT: Koivunen, Erkki I.  
; TITLE OF INVENTION: NOVEL INTEGRIN-BINDING PEPTIDES  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/212,186A  
; FILING DATE: 11-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9861  
; TELECOMMUNICATION INFORMATION:

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-212-186A-2

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
| | | | |  
Db 1 GRGDSP 6

## RESULT 8

US-38-169-743-4  
; Sequence 4, Application US/08169743  
; Patent No. 5547936  
; GENERAL INFORMATION:  
; APPLICANT: RUOSLAHTI, ERKKI I.  
; APPLICANT: PIERSCHBACHER, MICHAEL D.  
; APPLICANT: GEHLEN, KURT R.  
; TITLE OF INVENTION: INHIBITION OF CELL MIGRATION WITH  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 90071

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/169,743  
; FILING DATE:  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/773,106  
; FILING DATE: 08-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWN, THERESA A.  
; REGISTRATION NUMBER: 32,547  
; REFERENCE/DOCKET NUMBER: P319102  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-169-743-4

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
| | | | |  
Db 1 GRGDSP 6

## RESULT 9

US-08-262-315-1

; Sequence 1, Application US/08262315  
; Patent No. 5550131  
; GENERAL INFORMATION:  
; APPLICANT: SUGIHARA, Hiroasada  
; APPLICANT: TERASHITA, Zenichi  
; APPLICANT: FUKUSHI, Hideto  
; TITLE OF INVENTION: 2-PIPERAZINONE COMPOUNDS AND THEIR USE  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wegner, Cantor, Mueller & Player  
; STREET: 1233 20th Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-8218

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/262,315  
; FILING DATE: 17-JUN-1994  
; CLASSIFICATION: 544

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 146136/1993  
; FILING DATE: 17-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 254142/1993  
; FILING DATE: 12-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cantor, Herbert I.  
; REGISTRATION NUMBER: 24,392  
; REFERENCE/DOCKET NUMBER: P-8700-24315

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-0400  
; TELEFAX: (202) 835-0605  
; TELEX: 440706

INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-262-315-1

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
| | | | |  
Db 1 GRGDSP 6

## RESULT 10

US-08-425-238-7  
; Sequence 7, Application US/08425238  
; Patent No. 5627263  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Koivinen, Erkki  
; TITLE OF INVENTION: No. 5627263el Integrin-Binding Peptides  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/425,238  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/158,001  
;; FILING DATE: 24-NOV-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-LA 9775  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-425-238-7

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 11  
US-08-178-482-6  
; Sequence 6, Application US/08178482  
; Patent No. 5629294  
; GENERAL INFORMATION:  
; APPLICANT: DIZEREGA, GERE S  
; APPLICANT: RODGERS, KATHLEEN E  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING  
; TITLE OF INVENTION: ADHESION FORMATION  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBBINS, DALGARN, BERLINER & CARSON  
; STREET: 201 NORTH FIGUEROA STREET, FIFTH FLOOR  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90012-2628  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/178,482  
; FILING DATE: 06-JAN-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/789,211  
; FILING DATE: 07-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SPITALS, JOHN P  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-314  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 6 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-178-482-6

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 12  
US-08-264-759-3  
; Sequence 3, Application US/08264759  
; Patent No. 5656442  
; GENERAL INFORMATION:  
; APPLICANT: Ginsberg, Mark H  
; TITLE OF INVENTION: CHARACTERIZATION OF PLATELET AGGREGATION  
; TITLE OF INVENTION: DISORDERS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5656442th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/264,759  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/976,528  
; FILING DATE: 16-NOV-1992  
; APPLICATION NUMBER: US 07/614,723  
; FILING DATE: 15-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April C  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: SCRI239P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-08-264-759-3

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 13  
US-08-445-745-49

Sequence 49, Application US/08445745  
Patent No. 5672585  
GENERAL INFORMATION:  
APPLICANT: Pierschbacher, Michael D.  
APPLICANT: Cheng, Soat  
APPLICANT: Craig, William S.  
APPLICANT: Tschopp, Juerg F.  
TITLE OF INVENTION: Methods and Composition for Treating  
TITLE OF INVENTION: Thrombosis  
NUMBER OF SEQUENCES: 168  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,745  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/171,068  
FILING DATE: 20-DEC-1993  
APPLICATION NUMBER: US 08/079,441  
FILING DATE: 18-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/050,73614  
FILING DATE: 14-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/681,119  
FILING DATE: 05-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/506,444  
FILING DATE: 06-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9829  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-445-745-49

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6  
DB 1 GRGDSF 6

## RESULT 14

US-08-608-697-1  
Sequence 1, Application US/08608697  
Patent No. 5703081  
GENERAL INFORMATION:  
APPLICANT: Akio MIYAKE et al.  
TITLE OF INVENTION: QUINOLONECARBOXYLIC ACID DERIVATIVES, THEIR  
PRODUCTION AND USE  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/608,697  
FILING DATE: February 29, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/207,091  
FILING DATE: March 8, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-608-697-1

Query Match

100.0%; Score 34; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
|||  
Db 1 GRGDSP 6

RESULT 15  
US-08-596-116A-69  
; Sequence 69, Application US/08596116A  
; Patent No. 5721213  
; GENERAL INFORMATION:  
; APPLICANT: SATO, Yoshimi  
; APPLICANT: HAYASHI, Yoshio  
; APPLICANT: KATADA, Jun  
; TITLE OF INVENTION: No. 5721213el Peptides, Active as Inhibitors of  
; TITLE OF INVENTION: Platelet Aggregation  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kenyon & Kenyon  
; STREET: 1025 Connecticut Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3+ Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
; SOFTWARE: WordPerfect 6.1 for Windows  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/596,116A  
; FILING DATE: 30 JAN 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP94/01611  
; FILING DATE: 29 SEP 1994  
; APPLICATION NUMBER: JP 245541  
; FILING DATE: 30 SEP 1993  
; APPLICATION NUMBER: JP 50602  
; FILING DATE: 22 MAR 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Toffenetti, Judith L.  
; REGISTRATION NUMBER: 33,048  
; REFERENCE/DOCKET NUMBER: 2002/3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-1776  
; TELEFAX: 202-429-0796  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-596-116A-69

Query Match 100.0%; Score 34; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
|||  
Db 1 GRGDSP 6

Search completed: April 16, 2004, 07:31:40  
Job time : 20.35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2004, 07:13:22 ; Search time 19.35 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-8

Perfect score: 35

Sequence: 1 GRGTP 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgm2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgm2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgm2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgm2\_6/ptodata/2/iaa/6C.COMB.pep.\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	6	1	US-08-169-743-2
2	35	100.0	6	1	US-08-178-482-7
3	35	100.0	6	1	US-08-234-997-1
4	35	100.0	6	2	US-08-747-137-74
5	35	100.0	6	2	US-08-747-137-176
6	35	100.0	6	3	US-08-956-699-1
7	35	100.0	6	4	US-09-606-477-1
8	35	100.0	6	4	US-09-485-653-1
9	35	100.0	6	4	US-09-548-697A-2
10	35	100.0	6	5	PCT-US93-05640-45
11	35	100.0	9	1	US-08-445-745-114
12	35	100.0	9	4	US-08-456-466-92
13	35	100.0	9	4	US-08-445-638-114
14	35	100.0	10	1	US-08-445-745-50
15	35	100.0	10	1	US-08-445-745-51
16	35	100.0	10	1	US-08-445-745-102
17	35	100.0	10	1	US-08-445-745-103
18	35	100.0	10	1	US-08-445-745-104
19	35	100.0	10	1	US-08-445-745-105
20	35	100.0	10	1	US-08-445-745-106
21	35	100.0	10	1	US-08-445-745-110
22	35	100.0	10	1	US-08-445-745-117
23	35	100.0	10	4	US-08-456-466-24
24	35	100.0	10	4	US-08-456-466-80
25	35	100.0	10	4	US-08-456-466-81
26	35	100.0	10	4	US-08-456-466-82
27	35	100.0	10	4	US-08-456-466-83

28	35	100.0	10	4	US-08-456-466-84	Sequence 84, Appl
29	35	100.0	10	4	US-08-456-466-86	Sequence 88, Appl
30	35	100.0	10	4	US-08-456-466-95	Sequence 95, Appl
31	35	100.0	10	4	US-08-445-638-50	Sequence 50, Appl
32	35	100.0	10	4	US-08-445-638-51	Sequence 51, Appl
33	35	100.0	10	4	US-08-445-638-102	Sequence 102, App
34	35	100.0	10	4	US-08-445-638-103	Sequence 103, App
35	35	100.0	10	4	US-08-445-638-104	Sequence 104, App
36	35	100.0	10	4	US-08-445-638-105	Sequence 105, App
37	35	100.0	10	4	US-08-445-638-106	Sequence 106, App
38	35	100.0	10	4	US-08-445-638-110	Sequence 110, App
39	35	100.0	10	4	US-08-445-638-117	Sequence 117, App
40	35	100.0	127	4	US-09-489-039A-13293	Sequence 13293, A
41	35	100.0	493	3	US-08-999-774A-12	Sequence 12, Appl
42	35	100.0	1005	4	US-09-206-942-41	Sequence 41, Appl
43	35	100.0	1011	4	US-09-206-942-39	Sequence 39, Appl
44	35	100.0	1306	3	US-08-999-774A-13	Sequence 13, Appl
45	31	88.6	6	1	US-07-720-222-46	Sequence 46, Appl

## ALIGNMENTS

RESULT 1  
US-08-169-743-2  
; Sequence 2, Application US/08169743  
; Patent No. 5547936  
; GENERAL INFORMATION:  
; APPLICANT: RUOSLAHTI, ERKKI I.  
; APPLICANT: PIERSCHACHER, MICHAEL D.  
; APPLICANT: GEHLEN, KURT R.  
; TITLE OF INVENTION: INHIBITION OF CELL MIGRATION WITH  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/169,743  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/773,106  
; FILING DATE: 08-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWN, THERESA A.  
; REGISTRATION NUMBER: 32,547  
; REFERENCE/DOCKET NUMBER: P319102  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-169-743-2

Query Match 100.0%; Score 35; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGTP 6  
|||||

```

Db          1 GRGDTp 6

RESULT 2
US-08-178-482-7
; Sequence 7, Application US/38178482
; Patent No. 5629294
; GENERAL INFORMATION:
; APPLICANT: DIZEREGA, GERE S
; APPLICANT: RODGERS, KATHLEEN E
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING
; TITLE OF INVENTION: ADHESION FORMATION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBBINS, DALGARN, BERLINER & CARSON
; STREET: 201 NORTH FIGUEROA STREET, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90012-2628
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/178,482
; FILING DATE: 06-JAN-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,231
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SPITALS, JOHN P
; REGISTRATION NUMBER: 29,215
; REFERENCE/DOCKET NUMBER: 1923-314
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-178-482-7

Query Match          100.0%; Score 35; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY          1 GRGDTp 6
Db          1 GRGDTp 6

RESULT 3
US-08-234-997-1
; Sequence 1, Application US/38234997
; Patent No. 5683867
; GENERAL INFORMATION:
; APPLICANT: SUMEDHA JAYASENA
; APPLICANT: GREG BIESECKER
; APPLICANT: LARRY GOLD
; APPLICANT: DREW SMITH
; APPLICANT: GARY KIRSCHENHEUTER
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: BLENDED
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beaton & Swanson, P.C.
; STREET: 4582 South Ulster Street Parkway, #403
; CITY: Denver

```

```

; STATE: Colorado
; COUNTRY: USA
; ZIP: 80237
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/234,997
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: JUNE 10, 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: JUNE 11, 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: SEPTEMBER 8, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,935
; FILING DATE: SEPTEMBER 17, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,507
; FILING DATE: FEBRUARY 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BARRY J. SWANSON
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 850-9900
; TELEFAX: (303) 850-9401
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-234-997-1

Query Match          100.0%; Score 35; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 GRGDTp 6
Db          1 GRGDTp 6

RESULT 4
US-08-747-137-74
; Sequence 74, Application US/08747137
; Patent No. 5945033
; GENERAL INFORMATION:
; APPLICANT: YEN, Richard C.K.
; TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US 08/747,137  
FILING DATE: 12-NOV-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,546  
FILING DATE: 14-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/069,831  
FILING DATE: 01-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/959,560  
FILING DATE: 13-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/641,720  
FILING DATE: 15-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 016197-000840US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
US-08-747-137-74

Query Match 100.0%; Score 35; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTF 6  
Db 1 GRGDTF 6

RESULT 5  
US-08-747-137-176  
Sequence 176, Application US/08747137  
Patent No. 5945033  
GENERAL INFORMATION:  
APPLICANT: YEN, Richard C.K.  
TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR  
THERAPEUTIC AND DIAGNOSTIC USE  
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,137  
FILING DATE: 12-NOV-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,546  
FILING DATE: 14-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/069,831  
FILING DATE: 01-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/959,560  
FILING DATE: 13-OCT-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/641,720  
FILING DATE: 15-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 016197-000840US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
INFORMATION FOR SEQ ID NO: 176:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
US-08-747-137-176

Query Match 100.0%; Score 35; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTF 6  
Db 1 GRGDTF 6

RESULT 6  
US-08-956-699-1  
Sequence 1, Application US/08956699  
Patent No. 6083696  
GENERAL INFORMATION:  
APPLICANT: GREG BIESECKER  
APPLICANT: SUMEDHA D. JAYASENA  
APPLICANT: LARRY GOLD  
APPLICANT: DREW SMITH  
APPLICANT: GARY P. KIRSCHENHEUTER  
TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS  
TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: BLENDED  
TITLE OF INVENTION: SELEX  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 100  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,699  
FILING DATE: OCTOBER 23, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/234,997  
FILING DATE: APRIL 28, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/714,131  
FILING DATE: JUNE 10, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: JUNE 11, 1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/117,991  
FILING DATE: SEPTEMBER 8, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/123,935  
FILING DATE: SEPTEMBER 17, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/199,507  
FILING DATE: FEBRUARY 22, 1994



ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX15/C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-956-699-1

Query Match 100.0%; Score 35; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
Db 1 GRGDTP 6

RESULT 7  
US-09-606-477-1  
Sequence 1, Application US/09606477  
Patent No. 6465189  
GENERAL INFORMATION:  
APPLICANT: GREG RIESECKER  
SUMEDHA D. JAYASENA  
LARRY GOLD  
DREW SMITH  
GARY P. KIRSCHENHEUTER  
TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY  
EXPONENTIAL ENRICHMENT: BLENDED  
SELEX  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 1745 Shea Center Drive, Suite 330  
CITY: Highlands Ranch  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80129  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/606,477  
FILING DATE: 23-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/956,699  
FILING DATE: OCTOBER 23, 1997  
APPLICATION NUMBER: 08/234,997  
FILING DATE: APRIL 28, 1994  
APPLICATION NUMBER: 07/714,131  
FILING DATE: JUNE 10, 1991  
APPLICATION NUMBER: 07/536,428  
FILING DATE: JUNE 11, 1990  
APPLICATION NUMBER: 08/117,991  
FILING DATE: SEPTEMBER 8, 1993  
APPLICATION NUMBER: 08/123,935  
FILING DATE: SEPTEMBER 17, 1993  
APPLICATION NUMBER: 08/199,507  
FILING DATE: FEBRUARY 22, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX15/C-CON

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-606-477-1

Query Match 100.0%; Score 35; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
Db 1 GRGDTP 6

RESULT 8  
US-09-485-653-1  
Sequence 1, Application US/09485653  
Patent No. 6555321  
GENERAL INFORMATION:  
APPLICANT: Daniel, Thomas O.  
TITLE OF INVENTION: Methods for Determining Cell Responses  
FILE REFERENCE: 22000.0085U1  
CURRENT APPLICATION NUMBER: US/09/485,653  
CURRENT FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: PCT/US98/17157  
PRIOR FILING DATE: 1998-08-19  
PRIOR APPLICATION NUMBER: 60/056,164  
PRIOR FILING DATE: 1997-08-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence; No. 6555321e =  
OTHER INFORMATION: synthetic construct  
US-09-485-653-1

Query Match 100.0%; Score 35; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
Db 1 GRGDTP 6

RESULT 9  
US-09-548-697A-2  
Sequence 2, Application US/09548697A  
Patent No. 6586187  
GENERAL INFORMATION:  
APPLICANT: American Cyanamid Company  
APPLICANT: Gopalsamy, Ariamala  
APPLICANT: Yang, Hui Y  
TITLE OF INVENTION: Methods for Solid Phase Combinatorial Synthesis of Integrin Inhi  
FILE REFERENCE: AHP-98202  
CURRENT APPLICATION NUMBER: US/09/548,697A  
CURRENT FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: 60/240,952  
PRIOR FILING DATE: 1999-04-14  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 2
; LENGTH: 6
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-548-697A-2

Query Match      100.0%; Score 35; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6
Db 1 GRGDTTP 6

RESULT 10
PCT-US93-05640-45
; Sequence 45, Application PC/TUS9305640
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; TITLE OF INVENTION: Contraceptive Vaccine
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05640
; FILING DATE: 19930610
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/897,883
; FILING DATE: June 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: UCT90-01AA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9543
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-05640-45

Query Match      100.0%; Score 35; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6
Db 1 GRGDTTP 6

RESULT 11
US-08-445-745-114
; Sequence 114, Application US/08445745
; Patent No. 5672585
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; Lukeman, David S.
; Cheng, Soan
; Craig, William S.

; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,745
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,068
; FILING DATE: 28-DEC-1993
; APPLICATION NUMBER: US 08/079,441
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,73614
; FILING DATE: 14-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,119
; FILING DATE: 25-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/506,444
; FILING DATE: 08-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa = (Pmp)"
US-08-445-745-114

Query Match      100.0%; Score 35; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6
Db 2 GRGDTTP 7

RESULT 12
US-08-456-466-92
; Sequence 92, Application US/08456466
; Patent No. 6395873
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; Lukeman, David S.
; Cheng, Soan
; Craig, William S.
```

```

; Tschopp, Juerg F.
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR TREATING
; THROMBOSIS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,466
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa=Pmp"
; SEQUENCE DESCRIPTION: SEQ ID NO: 92:
US-08-456-466-92

Query Match 100.0%; Score 35; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTP 6
Db 2 GRGDTP 7

RESULT 13
US-08-445-638-114
; Sequence 114, Application US/08445638
; Patent No. 6521594
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; THROMBOSIS
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,638
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,068
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: US 08/079,441
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,73614
; FILING DATE: 14-APR-1993
; APPLICATION NUMBER: US 07/681,119
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/506,444
; FILING DATE: 06-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa = {Pmp}"
US-08-445-638-114

Query Match 100.0%; Score 35; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTP 6
Db 2 GRGDTP 7

RESULT 14
US-08-445-745-50
; Sequence 50, Application US/08445745
; Patent No. 5672585
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; THROMBOSIS
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/445,745
; FILING DATE:
; CLASSIFICATION: 514

```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/171,068  
;; FILING DATE: 20-DEC-1993  
;; APPLICATION NUMBER: US 08/079,441  
;; FILING DATE: 18-JUN-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/050,73614  
;; FILING DATE: 14-APR-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/681,119  
;; FILING DATE: 05-APR-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/506,444  
;; FILING DATE: 06-APR-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-LA 9829  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 2  
;; OTHER INFORMATION: /note= "Xaa = (Pen)"  
US-08-445-745-50

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred.No. 0.91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
Db 3 GRGDTP 8

RESULT 15  
US-08-445-745-51  
Sequence 51, Application US/08445745  
Patent No. 5672585  
GENERAL INFORMATION:  
APPLICANT: Pierschbacher, Michael D.  
APPLICANT: Cheng, Soan  
APPLICANT: Craig, William S.  
APPLICANT: Tschopp, Juerg F.  
TITLE OF INVENTION: Methods and Composition for Treating  
TITLE OF INVENTION: Thrombosis  
NUMBER OF SEQUENCES: 168  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,745  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/171,068  
FILING DATE: 20-DEC-1993  
APPLICATION NUMBER: US 08/079,441

;; FILING DATE: 18-JUN-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/050,73614  
;; FILING DATE: 14-APR-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/681,119  
;; FILING DATE: 05-APR-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/506,444  
;; FILING DATE: 06-APR-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-LA 9829  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 51:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 2  
;; OTHER INFORMATION: /note= "Xaa = (Pen)"  
US-08-445-745-51

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred.No. 0.91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
Db 3 GRGDTP 8

Search completed: April 16, 2004, 07:31:40  
Job time : 19.35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein: - protein search, using sw model

Run on: April 16, 2004, 06:55:57 ; Search time: 65.55 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588B-8

Perfect score: 35

Sequence: 1 GRGDTTP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 3

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	6	2 AAR36709	Aar36709 Adhesion
2	35	100.0	6	2 AAR47384	Aar47384 PH-30 bet
3	35	100.0	6	2 AAR94571	Aar94571 RGD pepti
4	35	100.0	6	2 AAW07430	Aaw07430 Synthetic
5	35	100.0	6	2 AAW84459	Aaw84459 RGD pepti
6	35	100.0	6	2 AAY50314	Aay50314 Neutroph
7	35	100.0	6	2 AAY31127	Aay31127 Non-cross
8	35	100.0	6	2 AAY31025	Aay31025 Non-cross
9	35	100.0	6	3 AAY80475	Aay80475 Cell adhe
10	35	100.0	6	3 AAB28163	Abp28163 Peptide #
11	35	100.0	6	4 AAB91972	Abp91972 Fibronect
12	35	100.0	6	5 ABB80076	Abp80076 Integrin
13	35	100.0	6	6 ADA14871	Ada14871 VNR-based
14	35	100.0	9	2 AAW15667	Aaw15667 Platelet
15	35	100.0	9	5 AAG78544	Aag78544 Integrin-
16	35	100.0	9	5 ABP61362	Abp61362 Anti-thro
17	35	100.0	10	2 AAW15663	Aaw15663 Platelet
18	35	100.0	10	2 AAW15656	Aaw15656 Platelet
19	35	100.0	10	2 AAW15657	Aaw15657 Platelet
20	35	100.0	10	2 AAW15599	Aaw15599 Platelet
21	35	100.0	10	2 AAW15655	Aaw15655 Platelet
22	35	100.0	10	2 AAW15658	Aaw15658 Platelet
23	35	100.0	10	2 AAW15659	Aaw15659 Platelet
24	35	100.0	10	2 AAW15670	Aaw15670 Platelet
25	35	100.0	10	5 ABP61353	Abp61353 Anti-thro

26	35	100.0	10	5 ABP61354	Abp61354 Anti-thro
27	35	100.0	10	5 ABP61351	Abp61351 Anti-thro
28	35	100.0	10	5 ABP61352	Abp61352 Anti-thro
29	35	100.0	10	5 ABP61365	Abp61365 Anti-thro
30	35	100.0	10	5 ABP61358	Abp61358 Anti-thro
31	35	100.0	10	5 ABP61294	Abp61294 Anti-thro
32	35	100.0	10	5 ABP61350	Abp61350 Anti-thro
33	35	100.0	493	4 AAE09330	Aae09330 Human int
34	35	100.0	897	6 ABB82571	Abb82571 H. influe
35	35	100.0	899	6 ABB82572	Abb82572 H. influe
36	35	100.0	1005	3 AAB01833	Aab01833 Haemophil
37	35	100.0	1011	3 AAB01832	Aab01832 Haemophil
38	35	100.0	1304	6 ABR64222	Abp64222 Angiogene
39	35	100.0	1304	6 ABO52995	Abo52995 Human spl
40	35	100.0	1306	4 AAE09331	Aae09331 Human int
41	32	91.4	455	3 AAG42811	Aag42811 Arabidops
42	32	91.4	455	3 AAG16791	Aag16791 Arabidops
43	32	91.4	458	3 AAG42810	Aag42810 Arabidops
44	32	91.4	458	3 AAG16790	Aag16790 Arabidops
45	32	91.4	1425	4 ABG23344	Abg23344 Novel hum

ALIGNMENTS

RESULT 1

AAR36709

ID AAR36709 standard; peptide; 6 AA.

AC AAR36709;

XX 25-MAR-2003 (revised)

DT 26-AUG-1993 (first entry)

XX Adhesion formation prevention RGD-contg. peptide.

XX Tissue repair; peritoneum; surgery; post-surgically; inhibition;

KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;

KW CNS; use.

OS Synthetic.

XX WO9308818-AL.

XX 13-MAY-1993.

XX 06-NOV-1992; 92WO-US009494.

XX 07-NOV-1991; 91US-00789231.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Dizerega GS, Rodgers KE;

XX WPI; 1993-167381/20.

XX Prevention of adhesion formation, partic. post-surgically - comprises

XX administering a RGD-contg. peptide for a time sufficient to permit tissue

XX repair.

XX Example; Page 18; 22pp; English.

XX The sequence is that of an RGD-contg. peptide which is used in a method

XX for prevention of adhesion formation for a time sufficient to permit

XX tissue repair. The method is used for minimizing or preventing adhesion

XX formation, partic. in the peritoneum following surgery, but also for e.g.

XX cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In

XX addn., the peptide inhibits platelet aggregation and does not induce

XX inflammation or trauma at the site of administration. (Updated on 25-MAR-

XX 2003 to correct PN field.)

XX Sequence 6 AA;

SQ

Fri Apr 16 11:04:27 2004

us-09-991-588b-8.rag

```

Query Match      100.0%; Score 35; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      1 GRGDTP 6

RESULT 2
AAR47384
ID AAR47384 standard; protein; 6 AA.
XX
AC AAR47384;
XX
DT 25-MAR-2003 (revised)
DT 22-JUN-1994 (first entry)
XX
DE PH-30 beta disintegrin control peptide.
XX
KW PH-20; PH-30; contraceptive; fertilisation; sperm surface protein;
KW vaccine; sperm-egg fusion.
XX
OS Rubella virus.
XX
PN WO9325233-A1.
XX
PD 23-DEC-1993.
XX
PF 10-JUN-1993; 93WO-US005640.
XX
PR 12-JUN-1992; 92US-00897883.
XX
PA (UYCO-) UNIV CONNECTICUT.
XX
PI Primakoff P, Myles DG;
XX
WPI; 1994-007200/01.
XX
Contraceptive vaccine for reducing sperm-egg fusion - comprises peptide
from sperm surface protein which stimulates antibody prodn.
XX
Example 7; Page 27; 79pp; English.
XX
Example 7 describes the use of PH-30 beta disintegrin peptides as
inhibitors of sperm fusion to egg plasma membrane. Modified peptides
AAR47382-83 and control peptides (AAR47384-85) were tested. From
CC observations it was concluded that the PH-30 beta disintegrin domain
CC represents an epitope which is critical in sperm-egg fusion. (Updated on
CC 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 6 AA;
Query Match      100.0%; Score 35; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      1 GRGDTP 6

RESULT 3
AAR94571
ID AAR94571 standard; peptide; 6 AA.
XX
AC AAR94571;
XX
DT 21-JUN-1996 (first entry)
XX
DE RGD peptide.
XX
KW Cytotactin; neuron; neurite; cell attachment; cell elongation; antibody.
(LJOL-) LA JOLLA CANCER RES FOUND.

Query Match      100.0%; Score 35; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      1 GRGDTP 6

RESULT 4
AAR07430
ID AAR07430 standard; peptide; 6 AA.
XX
AC AAR07430;
XX
DT 25-MAR-2003 (revised)
DT 21-JAN-1997 (first entry)
XX
DE Synthetic, preferred tumour invasion-inhibitory peptide.
XX
KW Tumour invasion; extracellular matrix; ECM; metastasis; RGD sequence;
cancer; inhibition; control.
XX
OS Synthetic.
XX
PN US5547936-A.
XX
PD 20-AUG-1996.
XX
PF 17-DEC-1993; 93US-00169743.
XX
PR 22-NOV-1983; 83US-00554821.
PR 10-MAR-1985; 85US-00744981.
PR 10-MAR-1988; 88US-00168530.
PR 09-SEP-1988; 88US-00242713.
PR 25-FEB-1991; 91US-00660526.
PR 10-APR-1991; 91US-00683585.
PR 08-OCT-1991; 91US-00773106.
PR 19-JUN-1992; 92US-00902742.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.

```

XX PI Getlsen KR, Pierschbacher XD, Ruoslahti EI;  
 XX DR WPI; 1996-392651/39.  
 XX XX Inhibiting tumour cell invasion through an extracellular matrix - using  
 PT peptide confg. the RGD sequence, partic. for preventing tumour  
 PT metastasis.  
 XX PS Claim: 2; Col: 7-8; 8pp; English.  
 XX CC AAW07430 is a preferred peptide identified in an assay for testing  
 CC peptides for tumour-invasion inhibitory activity. The peptides (contg.  
 CC the RGD sequence) shows significant inhibition of tumour invasion of the  
 CC ECM. Other peptides tested (see AAW07431-W06433) did show inhibitory  
 CC activity but to a lesser extent than peptide AAW07430. The peptides  
 CC identified can be used to treat cancer and to prevent metastasis, in  
 CC partic. invasion of the extracellular matrix (ECM). The peptides are  
 CC soluble. (Updated on 25-MAR-2003 to correct PF field.)  
 XX CC  
 XX SQ Sequence 6 AA;  
 Query Match 100.0%; Score 35; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 1 GRGDTTP 6  
 DB 1 GRGDTTP 6  
 RESULT 5  
 ID AAW84459 standard; peptide; 6 AA.  
 XX AC AAW84459;  
 XX DT 25-MAR-1999 (first entry)  
 XX DE RGD peptide that stimulates cell invasion by S. pyogenes 90-226.  
 XX DE Streptococcus pyogenes 90-226; pathogen; inhibition; cell adhesion;  
 XX KW cell invasion; treatment; bacterial infection; fungal infection.  
 XX OS Synthetic.  
 XX PN W09856408-A2.  
 XX PD 17-DEC-1998.  
 XX PF 10-JUN-1998; 98WO-US012019.  
 XX PR 10-JUN-1997; 97US-0049124P.  
 XX PA (MNU ) UNIV MINNESOTA.  
 XX PI Cleary PP, Cue DR;  
 XX DR WPI; 1999-080856/07.  
 XX PT Method for treating mammal infected by pathogenic microorganism -  
 PT comprises administering to mammal composition comprising inhibitory  
 PT compound which inhibits adherence to or invasion of a cell by  
 PT microorganism.  
 XX PS Example 2; Page 43; 88pp; English.  
 XX CC The present sequence represents a RGD peptide that is able to stimulate  
 CC cell invasion by Streptococcus pyogenes 90-226. The peptides were used in  
 CC the course of the invention. The specification describes the treatment of  
 CC a mammal infected by a pathogenic microorganism which comprises  
 CC administering an inhibitory compound to inhibit adherence to or invasion  
 CC of cells by the pathogen. The method is used in the treatment of

CC bacterial or fungal infection  
 XX Sequence 6 AA;  
 SQ Query Match 100.0%; Score 35; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 1 GRGDTTP 6  
 DB 1 GRGDTTP 6  
 RESULT 6  
 ID AAY50314 standard; peptide; 6 AA.  
 XX AC AAY50314;  
 XX DT 12-JAN-2000 (first entry)  
 XX DE Neutrophil-activating pancreatic derived peptide 114.  
 XX KW Cell activation; pancreas; treatment; cardiovascular disease; trauma;  
 KW inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;  
 KW organ rejection; ischemia; Alzheimer's disease; myocardial infarction;  
 KW haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;  
 KW trauma; protease inhibitor; hypertension; sepsis.  
 XX OS Unidentified.  
 XX PN W09946367-A2.  
 XX PD 16-SEP-1999.  
 XX PR 11-MAR-1999; 99WO-US005247.  
 XX PR 11-MAR-1999; 98US-00038894.  
 XX PA (CELL-) CELL ACTIVATION INC.  
 XX PA (REGC ) UNIV CALIFORNIA.  
 XX PA (SCRI ) SCRIPPS RES INST.  
 XX PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;  
 XX DR WPI; 1999-580234/49.  
 XX PT Use of cell activating compositions in developing products for diagnosis  
 PT and treatment of e.g. cardiovascular, inflammatory, autoimmune or  
 PT Alzheimer's disease, trauma, arthritis, organ rejection, diabetes, stroke  
 PT or ischemia.  
 XX PS Example 9; Page 184; 184pp; English.  
 XX CC This invention describes a novel method for the use and preparation of  
 CC cell activating compositions which involves preparing a cell activating  
 CC composition comprising (a) homogenizing pancreatic tissue in buffer at  
 CC about neutral or higher pH to produce a homogenate; (b) removing  
 CC particulates from the homogenate; (c) optionally incubating the resulting  
 CC homogenate, with particulates removed, with a protease; and (d)  
 CC fractionating the homogenate and selecting fractions that exhibit cell  
 CC activation activity. The methods can be used for improving treatment  
 CC outcome or reducing risk of treatment of e.g. cardiovascular disease,  
 CC inflammatory disease, trauma, autoimmune diseases, arthritis, organ  
 CC rejection, diabetes and diabetic complications, stroke, ischemia,  
 CC Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic  
 CC retinopathy, diabetes, venous insufficiency, unstable angina or trauma.  
 CC They can be used in the veterinary treatment of a non-human subject.  
 CC Protease inhibitors can be used to lower cell activation resulting from  
 CC these diseases and deficiencies. The detection of an elevated level of  
 CC hydrogen peroxide can be used to detect an inflammatory condition. An  
 CC elevated level of hydrogen peroxide in plasma or whole blood and in the  
 CC presence of superoxide dismutase (SOD) indicates leukocyte up regulation,

CC e.g. indicative of the onset of an acute cardiovascular disorders, such  
 CC as disease onset or ischemic complications. An elevated level of hydrogen  
 CC peroxide in plasma or whole blood and a low level in the presence of SOD  
 CC is indicative of a chronic or immune compromised condition e.g.  
 CC hypertension or sepsis. AAY50201-Y50334 represent peptides used in the  
 CC method of the invention  
 XX  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 35; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTP 6  
 Db 1 GRGDTP 6

## RESULT 7

AAY31127  
 ID AAY31127 standard; peptide; 6 AA.

XX AC AAY31127;

XX DT 21-OCT-1999 (first entry)

XX DE Non-crosslinked protein particle peptide 176.

XX KW Non-crosslinked protein particle; diagnostic; therapy; monodisperse;  
 XX albumin; haemoglobin; nanometer; micrometer; clearance.

XX OS Synthetic.

XX PN US5945033-A.

XX PD 31-AUG-1999.

XX PF 12-NOV-1996; 96US-00747137.

XX PR 15-JAN-1991; 91US-00641720.

XX PR 13-OCT-1992; 92US-00959560.

XX PR 01-JUN-1993; 93US-00069831.

XX PR 14-MAR-1994; 94US-00212546.

XX PA (HEMO-) HEMOSPHERE INC.

XX PI Yen RCK;

XX DR WPI; 1999-508153/42.

XX PT Non-crosslinked protein particles for therapeutic and diagnostic use.

XX PS Example 22; Col 123-124; 65pp; English.

XX CC This invention describes a novel aqueous suspension of monodisperse  
 CC particles on non-crosslinked, non-denatured albumin (50-5000 nm) which is  
 CC stable against dissolving upon dilution with an alcohol-free aqueous  
 CC medium. The method involves (a) forming an aqueous solution containing  
 CC albumin and hemoglobin and (b) treating the aqueous solution with an  
 CC alcohol to cause the solution to become turbid. The particles are useful  
 CC as agents for in vivo administration, either of their own administration  
 CC or as a vehicle for other therapeutic or diagnostic agents. The method  
 CC permits the formation of albumin and hemoglobin particles in the  
 CC nanometer and micrometer size range, in a form closer to their natural  
 CC form than the forms of the prior art. The particles therefore constitute  
 CC a more closely controlled agent for in vivo administration, with greater  
 CC ease of clearance from the body after their period of usefulness.  
 CC AAY30952-Y31135 represent peptides used in the method of the invention

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 35; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GRGDTP 6  
 Db 1 GRGDTP 6

## RESULT 8

AAY31025  
 ID AAY31025 standard; peptide; 6 AA.

XX AC AAY31025;

XX DT 21-OCT-1999 (first entry)

XX DE Non-crosslinked protein particle peptide 74.

XX KW Non-crosslinked protein particle; diagnostic; therapy; monodisperse;  
 XX albumin; haemoglobin; nanometer; micrometer; clearance.

XX OS Synthetic.

XX PN US5945033-A.

XX PD 31-AUG-1999.

XX PF 12-NOV-1996; 96US-00747137.

XX PR 15-JAN-1991; 91US-00641720.

XX PR 13-OCT-1992; 92US-00959560.

XX PR 01-JUN-1993; 93US-00069831.

XX PR 14-MAR-1994; 94US-00212546.

XX PA (HEMO-) HEMOSPHERE INC.

XX PI Yen RCK;

XX DR WPI; 1999-508153/42.

XX PT Non-crosslinked protein particles for therapeutic and diagnostic use.

XX PS Example 22; Col 81-82; 65pp; English.

XX CC This invention describes a novel aqueous suspension of monodisperse  
 CC particles on non-crosslinked, non-denatured albumin (50-5000 nm) which is  
 CC stable against dissolving upon dilution with an alcohol-free aqueous  
 CC medium. The method involves (a) forming an aqueous solution containing  
 CC albumin and hemoglobin and (b) treating the aqueous solution with an  
 CC alcohol to cause the solution to become turbid. The particles are useful  
 CC as agents for in vivo administration, either of their own administration  
 CC or as a vehicle for other therapeutic or diagnostic agents. The method  
 CC permits the formation of albumin and hemoglobin particles in the  
 CC nanometer and micrometer size range, in a form closer to their natural  
 CC form than the forms of the prior art. The particles therefore constitute  
 CC a more closely controlled agent for in vivo administration, with greater  
 CC ease of clearance from the body after their period of usefulness.  
 CC AAY30952-Y31135 represent peptides used in the method of the invention

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 35; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTP 6  
 Db 1 GRGDTP 6

## RESULT 9

AAY80475  
 ID AAY80475 standard; peptide; 6 AA.

XX



AC AAY80475;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Cell adhesion peptide #10.  
 XX  
 KW Bone regenerative; osteopathic; osseous tissue; reconstitution;  
 KW scaffold matrix; bone formation promoter; bone resorption inhibitor;  
 KW cell adhesion; osteoblast; osteoclast; bone defect; fracture.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200004941-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 22-JUL-1999; 99WO-US016800.  
 XX  
 PR 24-JUL-1998; 98US-0012348.  
 XX  
 PA (PHAR-) PHARMACAL BIOTECHNOLOGIES INC.  
 XX  
 PI Budny JA;  
 XX  
 DR WPI; 2000-195084/17.  
 XX  
 PT System for reconstructing osseous tissue, useful e.g. for treating  
 PT fractures, comprises scaffold containing promoter of bone formation and  
 PT inhibitor of bone resorption.  
 XX  
 PS Claim 14; Page 31; 44pp; English.  
 XX  
 CC The invention relates to a novel system for reconstitution of osseous  
 CC tissue comprising a scaffold carrying a compound (I) that promotes bone  
 CC formation and a component that decreases bone resorption (II). (I)  
 CC induces migration and adhesion of osteoblasts and osteoclasts and (II)  
 CC inhibits proteolysis (specifically by plasmin) of extracellular matrix.  
 CC (I) is preferably selected from: selectin or selectin binding fragments,  
 CC proteins, and peptides that facilitate cell adhesion, plasminogen  
 CC activator inhibitors, protease inhibitors and metalloprotease inhibitors.  
 CC The peptides AAV80466-Y80492 are claimed examples of cell adhesion  
 CC peptides used in the system of the invention. The system is used to  
 CC replace, remodel or correct bone defects, e.g. fractures, fissures or  
 CC bone mass loss. Incorporation of (I) into the scaffold results in rapid  
 CC seeding by osteoblasts and the development of an organic matrix, i.e. the  
 CC preformed scaffold replaces the rate-determining step of extracellular  
 CC matrix formation. The scaffold can be designed to have a predetermined  
 CC resorption/degradation rate, and may include regulatory compounds for  
 CC specific cell types  
 XX  
 SQ Sequence 6 AA;  
 Query Match 100.0%; Score 35; DB 3; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDTP 6  
 Db |||||  
 1 GRGDTP 6  
 RESULT 10  
 AAB28163  
 ID AAB28163 standard; peptide; 6 AA.  
 XX  
 AC AAB28163;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Peptide #3 used to assay vitronectin receptor binding inhibition.  
 XX  
 KW Antiviral; antiinflammatory; cyrostatic; vasotropic; antirheumatic;  
 KW antiarthritic; ophthalmological; osteopathic;

KW amido-carboxylic acid derivative; cancer; angiogenesis;  
 KW neovascularisation; macular degeneration; glaucoma;  
 KW blindness rheumatoid arthritis; restenosis; viral infection;  
 KW bone resorption; osteoporosis; osteopenia; periodontal disease;  
 KW hyperparathyroidism; Paget's disease; integrin vitronectin receptor;  
 KW alpha\_vbeta\_3; melanoma.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2000061545-A1.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 13-APR-2000; 2000WO-US010027.  
 XX  
 PR 14-APR-1999; 99US-00291470.  
 XX  
 PA (AMHP ) AMERICAN HOME PROD CORP.  
 XX  
 PI Gopalsamy A, Yang HY;  
 XX  
 DR WPI; 2000-687029/67.  
 XX  
 PT New amido-carboxylic acid derivatives are integrin inhibitors used for  
 PT treating e.g. cancer, restenosis, osteoporosis, viral infection and bone  
 PT disease.  
 XX  
 PS Disclosure; Page 17; 67pp; English.  
 XX  
 CC The present invention relates to amido-carboxylic acid derivatives. The  
 CC amido-carboxylic acid derivatives may be used for treating cancer,  
 CC angiogenesis, neovascularisation, macular degeneration, glaucoma,  
 CC blindness, rheumatoid arthritis, restenosis, smooth cell proliferation  
 CC and migration, vascular endothelial cell proliferation and migration,  
 CC viral infection (characterised by bone resorption of mineralised  
 CC tissues), osteoporosis, hypercalcaemia of malignancy, osteopenia due to  
 CC bone metastasis, periodontal disease, hyperparathyroidism, periarticular  
 CC erosions in rheumatoid arthritis, Paget's disease, immobilisation-induced  
 CC osteopenia or glucocorticoid treatment, or diseases characterised by some  
 CC resorption of mineralised tissues. The present sequence is a peptide,  
 CC which was used as a reference compound in an assay for the ability of the  
 CC amido-carboxylic acid derivatives of the present invention to inhibit  
 CC integrin vitronectin receptor (alpha\_vbeta\_3) binding. Integrin  
 CC alpha\_vbeta\_3 has been shown to mediate the invasion of cancerous  
 CC melanoma cells into healthy tissue and to protect these cells against  
 CC apoptosis. In addition, vitronectin receptor antagonists have been shown  
 CC to inhibit the growth of various solid tumours of human origin  
 XX  
 SQ Sequence 6 AA;  
 Query Match 100.0%; Score 35; DB 3; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDTP 6  
 Db |||||  
 1 GRGDTP 6  
 RESULT 11  
 AAB91972  
 ID AAB91972 standard; peptide; 6 AA.  
 XX  
 AC AAB91972;  
 XX  
 DT 22-JUN-2001 (first entry)  
 XX  
 DE Fibronectin fragment and fibrin related peptide SEQ ID NO:1148.  
 XX  
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidy; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Homo sapiens.  
OS Synthetic.  
XX WO200069900-A2.  
XX 23-NOV-2000.  
XX 17-MAY-2000; 2000WO-US013576.  
XX 17-MAY-1999; 99US-0134406P.  
XX 10-SEP-1999; 99US-0153406P.  
XX 15-OCT-1999; 99US-0159783P.  
XX (CONJ-) CONJUCHEM INC.  
XX Bridon DP, Ezrin AV, Milner PG, Holmes DL, Thibaudau K;  
XX WPI; 2001-112059/12.  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity.  
XX Disclosure; Page 571; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide) and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity in  
CC vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention  
XX  
XX Sequence 6 AA;  
SQ  
Query Match 100.0%; Score 35; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDTP 6  
DB |||||  
1 GRGDTP 6  
RESULT 12  
AAB80076  
ID ABB80076 standard; peptide; 6 AA.  
XX  
XX ABB80076;  
XX 27-AUG-2002 (first entry)  
XX Integrin modulator peptide #3.  
XX Neurodegenerative disease; amyloid; neuroprotective; anti-Alzheimer's;  
XX Alzheimer's disease; integrin; therapeutic agent.  
XX Unidentified.  
XX WO200226107-A2.  
XX 04-APR-2002.  
XX 25-SEP-2001; 2001WO-US023788.  
XX

XX 25-SEP-2000; 2000US-0235374P.  
XX (REGC ) UNIV CALIFORNIA.  
XX Lynch G, Xiaoning B, Gall CM;  
XX WPI; 2002-416420/44.  
XX Experimental models for studying neurodegenerative diseases associated  
PT with amyloid accumulation, useful for studying e.g. Alzheimer's disease  
PT and for identifying potential therapeutic agents.  
XX Claim 15; Page 63; 80pp; English.  
XX The invention relates to experimental models for studying  
CC neurodegenerative diseases associated with amyloid accumulation. The  
CC activity of peptides of the invention may be described as neuroprotective  
CC and anti-Alzheimer's. They act by inhibiting amyloid accumulation. The  
CC experimental model may be used to study neurodegenerative diseases  
CC associated with amyloid accumulations, e.g. Alzheimer's disease and to  
CC identify potential therapeutic agents for treating those diseases. The  
CC current sequence represents an integrin modulator peptide of the  
CC invention  
XX  
XX Sequence 5 AA;  
SQ  
Query Match 100.0%; Score 35; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDTP 6  
DB |||||  
1 GRGDTP 6  
RESULT 13  
ADA14871  
ID ADA14871 standard; peptide; 6 AA.  
XX  
XX ADA14871;  
XX 06-NOV-2003 (first entry)  
XX VnR-based adhesion inhibiting RGD-containing peptide #1.  
XX gene therapy; interphotoreceptor matrix component; IPMC; ocular disorder;  
XX macular degeneration; photoreceptor death; retinal detachment;  
XX VnR-based adhesion.  
XX Synthetic.  
XX US2002160954-A1.  
XX 31-OCT-2002.  
XX 08-NOV-2001; 2001US-00007270.  
XX 29-OCT-1998; 98US-00183972.  
XX 29-OCT-1999; 99US-00430195.  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX Hageman GS, Kuehn MH;  
XX WPI; 2003-238235/23.  
XX New isolated or recombinant interphotoreceptor matrix component  
PT polynucleotide and polypeptide, useful for diagnosing, preventing,  
PT treating or prognosticating ocular disorders, e.g. macular degeneration  
PT or retinal detachment.  
XX Example 7; Page 69; 76pp; English.  
PS

XX The invention relates to an isolated or recombinant interphotoreceptor  
 CC matrix component (IPMC) polynucleotide. Also disclosed is a vector  
 CC comprising a promoter of an interphotoreceptor matrix component (IPMC)  
 CC gene operatively linked to the IPMC polynucleotide. The IPMC  
 CC polynucleotides, polypeptides and antibodies are useful for diagnosing,  
 CC preventing, treating or prognosticating ocular disorders, e.g. macular  
 CC degeneration, photoreceptor death or retinal detachment. They are also  
 CC useful for identifying a compound capable of modulating IPMC gene  
 CC expression in a cell. The present sequence represents the amino acid  
 CC sequence of the Vnr-based adhesion inhibiting RGD-containing peptide #1.  
 XX  
 SQ Sequence 6 AA;  
 Query Match 100.0%; Score 35; DB 6; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDTP 6  
 DB 1 GRGDTP 6  
 RESULT 14  
 AAW15667  
 ID AAW15667 standard; peptide; 9 AA.  
 AC AAW15667;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 11-JUN-1997 (first entry)  
 DT  
 XX Platelet aggregation inhibitor #92.  
 DE  
 XX Platelet aggregation inhibitor; RGD analogue; cyclic peptide; fibrinogen;  
 KW hydrophobically enhanced analogue; blood platelet; endothelial surface;  
 KW blood vessel; serum protein; GP IIb/IIIa glycoprotein complex; integrin;  
 KW plasma membrane; thrombosis; cell adhesion receptor; fibronectin;  
 KW vitronectin receptor; vascular graft occlusion; therapy.  
 XX  
 OS Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 PH Modified-site 1..8  
 FT /note= "forms peptide bond to create cyclic peptide"  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "beta1, beta-pentamethylene-beta-mercaptopropionic  
 FT acid"  
 FT  
 XX US5612311-A.  
 PN 18-MAR-1997.  
 PD  
 XX 22-DEC-1994; 94US-00363963.  
 XX  
 XX 06-APR-1990; 90US-00506444.  
 PR 05-APR-1991; 91US-00681119.  
 PR 14-APR-1993; 93US-00050736.  
 PR 02-MAR-1994; 94US-00204817.  
 XX  
 XX (LJOL-) LA JOLLA CANCER RES FOUND.  
 PA  
 XX Craig WS, Lukeman DS, Cheng S, Tschopp JF, Pierschbacher MD;  
 PI WPI; 1997-192139/17.  
 XX  
 XX RGD-contg. peptide(s) that inhibit platelet aggregation - useful for  
 PT treating thrombosis.  
 PT  
 XX Example 5; Col 75; 50pp; English.  
 PS  
 XX AAW15576-W15695 represent platelet aggregation inhibitors. All of these  
 CC

CC sequences are hydrophobically enhanced RGD peptide analogues. The  
 CC interaction of blood platelets with the endothelial surface of injured  
 CC blood vessels and with other platelets (platelet aggregation) is a major  
 CC factor in the course of development of thrombi. Thrombosis is a serious  
 CC condition which can cause tissue damage and eventually death (if  
 CC untreated). Platelet aggregation is dependent upon the binding of  
 CC fibrinogen and other serum proteins to the GP IIb/IIIa glycoprotein  
 CC complex on the platelet plasma membrane. GP IIb/IIIa is a member of the  
 CC integrin family of cell adhesion receptors, which are known to recognise  
 CC a RGD tripeptide recognition sequence. The peptides inhibit platelet  
 CC aggregation without prolonging bleeding time. These sequences have high  
 CC affinity for the IIb/IIIa receptor and low affinity for the fibronectin  
 CC and vitronectin receptors. The peptides are used as platelet aggregation  
 CC inhibitors for treating thrombosis and vascular graft occlusion. (Updated  
 CC on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 35; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDTP 6  
 DB 2 GRGDTP 7  
 RESULT 15  
 AAG78544  
 ID AAG78544 standard; peptide; 9 AA.  
 XX  
 AC AAG78544;  
 XX  
 XX 04-MAR-2002 (first entry)  
 DT Integrin-binding peptide.  
 DE  
 XX Embolisation therapy; angiogenic-dependant disease; cancer;  
 KW drug delivery; vaccine; microsphere; precancerous; therapeutic factor;  
 KW integrin-binding; cytostatic.  
 XX  
 OS Unidentified.  
 OS  
 XX WO200172281-A2.  
 PN  
 XX 04-OCT-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009619.  
 PF  
 XX 24-MAR-2000; 2000US-0191899P.  
 PR  
 XX (BIOS-) BIOSPHERE MEDICAL INC.  
 PA  
 XX Vogel J, Boschetti E;  
 PI WPI; 2002-034141/04.  
 XX  
 XX Microspheres for active embolization comprising a crosslinked hydrophilic  
 FT polymer and a drug or vaccine are useful for site-directed therapy via  
 FT the blood supply, particularly of tumors.  
 FT  
 XX Disclosure; Page 34; 73pp; English.  
 PS  
 XX The invention relates to a microsphere suitable for active embolisation,  
 CC comprising a biocompatible, cross-linked and substantially hydrophilic  
 CC polymer and a drug or vaccine, the activity of which may be described as  
 CC cytostatic. Compositions and methods of the invention may be used in the  
 CC treatment of angiogenic-dependant diseases including cancer and  
 CC precancerous disorders. The microspheres are used for site-directed  
 CC delivery of therapeutic drugs, polynucleotides or vaccines, particularly  
 CC for treatment of tumours. The current sequence represents an integrin  
 CC binding peptide that may be used as a transfection agent to increase the  
 CC efficiency of transfer of a bioactive therapeutic factor into cells

XX  
SQ Sequence 9 AA;

Query Match 100.0%; Score 35; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGTP 6  
    |||||  
Db 4 GRGTP 9

Search completed: April 16, 2004, 07:20:54  
Job time : 66.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein: - protein search, using sw model

Ran on: April 16, 2004, 06:55:57 ; Search time: 76.475 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588b-10

Perfect score: 43

Sequence: 1 GRGDSPC 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 8%

Maximum Match: 100%  
Listing first 45 summaries

Database : A: Genesep 25Jan04: \*  
1: genesep1980s: \*  
2: genesep1990s: \*  
3: genesep2000s: \*  
4: genesep2001s: \*  
5: genesep2002s: \*  
6: genesep2003as: \*  
7: genesep2003bs: \*  
8: genesep2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	7	AAR11505	Cell att
2	43	100.0	7	AAR36716	Adhesion
3	43	100.0	7	AAR25183	RGD-pepti
4	43	100.0	7	AAR18801	RGD-pepti
5	43	100.0	7	AAR37767	Receptor
6	43	100.0	7	AAY43230	RGD-Conta
7	43	100.0	7	AAY31133	Non-cross
8	43	100.0	7	AAY80477	Cell adhe
9	43	100.0	7	AAB91971	Fibronect
10	43	100.0	7	AAB28378	Fibronect
11	43	100.0	7	ABB09833	Amino aci
12	43	100.0	8	AAR21014	Cyclised
13	43	100.0	8	AAR21005	Cyclised
14	43	100.0	8	AAR48503	Integrin
15	43	100.0	8	AAR48495	Integrin
16	43	100.0	8	AAB01878	Collagen
17	43	100.0	8	AAR19962	RNA inter
18	43	100.0	8	ABR81895	Integrin
19	43	100.0	9	AAR15666	Platelet
20	43	100.0	9	AAB91979	Fibronect
21	43	100.0	9	ABP61361	Anti-thro
22	43	100.0	9	AAO17746	Conformat
23	43	100.0	9	AAU97521	Synthetic
24	43	100.0	10	AAR36715	Adhesion
25	43	100.0	10	AAR72349	Cyclic in

## ALIGNMENTS

### RESULT 1

ID	AAR11505	standard; peptide; 7 AA.
XX	XX	
AC	AAR11505;	
DT	12-JUN-1991	(first entry)
XX	XX	
DE	Cell attachment promoting peptide.	
XX	XX	
KW	Fibrin; aggregation.	
XX	XX	
OS	Synthetic.	
XX	XX	
PH	Key	Location/Qualifiers
FT	Active-site	2. .4
XX	XX	
PN	US4988621-A.	
XX	XX	
PD	29-JAN-1991.	
XX	XX	
PF	10-DEC-1987;	87US-00131130.
XX	XX	
PR	24-MAY-1985;	85US-00738078.
XX	XX	
PI	{JOLLE-} LA JOLLA CANCER FOU.	
XX	XX	
PI	Ruoslahti EI, Hayman EG, Pierschbacher MD;	
XX	XX	
DR	WPI; 1991-116404/16.	
XX	XX	
PT	Peptide(s) contg. arginine-glycine-aspartic acid sequence - used to prevent and reverse cell attachment or to promote cell aggregation.	
XX	XX	
PS	Disclosure; Page 8; 12pp; English.	
XX	XX	
CC	The peptide, or shorter versions contg. the RGD active site from fibronectin, can be used to prevent and reverse attachment of cells to substrates. This can be used in cell prodn., fermentation, cell line prep., cell matrix prodn., diagnostics and therapy. The peptide can be used for eg mobilisation of bone marrow cells; prevention and reversal of attachment of disseminated tumour cells locally such as in the case of an operation performed in the peri-toneal cavity, to prevent adhesions and scar formations locally as in the case of eye operations, for prophylactic inhibition of E. coli binding to epithelial cells of the urinary tract or intestine, diagnosis and treatment of E. coli related infections, and identification of various pathogenic bacterial strains. The peptide is pref. prepd. by solid phase synthesis. See also AAR11506	

26	43	100.0	10	2	AAR07433	Synthetic
27	43	100.0	10	2	AAR15594	Platelet
28	43	100.0	10	2	AAR15593	Platelet
29	43	100.0	10	2	AAR65002	Synthetic
30	43	100.0	10	2	AAR68564	Cyclic RG
31	43	100.0	10	2	AAR51224	Integrin
32	43	100.0	10	2	AAR66860	Peptide u
33	43	100.0	10	2	AAR97029	Peptide u
34	43	100.0	10	2	AAR03843	RGD motif
35	43	100.0	10	3	AAR54910	RGD-conta
36	43	100.0	10	3	AAR32500	Integrin
37	43	100.0	10	3	AAR52146	RGD conta
38	43	100.0	10	3	AAR28161	Peptide #
39	43	100.0	10	5	ABR08334	RGD cycli
40	43	100.0	10	5	ABR61288	Anti-thro
41	43	100.0	10	5	ABP61289	Conformat
42	43	100.0	10	5	AAO17737	Synthetic
43	43	100.0	10	5	AAU97511	Synthetic
44	43	100.0	10	5	AAU97518	Synthetic
45	43	100.0	10	6	ADA14874	VNR-based

```

XX SQ Sequence 7 AA;
Query Match 100.0%; Score 43; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 2
AAR36716
ID AAR36716 standard; peptide; 7 AA.
XX
XX AAR36716;
XX
XX 25-MAR-2003 (revised)
DT 26-AUG-1993 (first entry)
XX
XX Adhesion formation prevention RGD-contg. peptide.
XX
XX Tissue repair; peritoneum; surgery; post-surgically; inhibition;
KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;
KW CNS; use.
XX
XX Synthetic.
XX
XX WO9308818-A1.
XX
XX 13-MAY-1993.
XX
XX 06-NOV-1992; 92WO-US009494.
XX
XX 07-NOV-1991; 91US-00789231.
XX
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
XX Dizerega GS, Rodgers KR;
XX
XX WPI; 1993-167381/20.
XX
XX Prevention of adhesion formation, partic. post-surgically - comprises
PT administering a RGD-contg. peptide for a time sufficient to permit tissue
PT repair.
XX
XX Example; Page 18; 22pp; English.
XX
XX The sequence is that of an RGD-contg. peptide which is used in a method
CC for prevention of adhesion formation for a time sufficient to permit
CC tissue repair. The method is used for minimising or preventing adhesion
CC formation, partic. in the peritoneum following surgery, but also for e.g.
CC cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In
CC addn., the peptide inhibits platelet aggregation and does not induce
CC inflammation or trauma at the site of administration. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
XX Sequence 7 AA;
Query Match 100.0%; Score 43; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 3
AAW25183
ID AAW25183 standard; peptide; 7 AA.
XX

```

```

AC AAW25183;
XX
DT 05-JAN-1998 (first entry)
XX
DE RGD-peptide capable of binding cell adhesion molecules.
XX
XX RGD; arginine; glycine; aspartic acid; cell adhesion molecule; binding;
KW bladder irrigation; tumour removal; endoscopic operation;
KW transurethral resection; cancer; neoplasia.
XX
XX Synthetic.
XX
XX DE19529939-A1.
XX
XX 20-FEB-1997.
XX
XX 15-AUG-1995; 95DE-01029909.
XX
XX 15-AUG-1995; 95DE-01029909.
XX
XX (FREP ) FRESSENIUS AG.
XX
XX Boehle A;
XX
XX WPI; 1997-133793/13.
XX
XX Endoscopic irrigation solns. - contg. peptide(s) that bind to cell
PT adhesion molecules.
XX
XX Claim 5; Page 8; 8pp; German.
XX
XX AAW25173-W25186 are peptides containing an RGD sequence or equivalent.
CC The peptides are capable of binding to cell adhesion molecules and are
CC used in aqueous irrigation solutions for use during and after endoscopic
CC operations. Preferred irrigation solutions are electrolyte-free and
CC contain 1 microg/ml to 100 mg/ml of one or more oligopeptides containing
CC the amino acid sequences: RGD, LDV, IDA, DGEA, GPRP, VTL, YIGSR, KOAGDV
CC and/or RBDV (given in one letter amino acid code). The solutions are
CC especially used for irrigating the bladder during and after tumour
CC removal by transurethral resection. The peptides protect against
CC recurrence of tumours
XX
XX Sequence 7 AA;
Query Match 100.0%; Score 43; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 4
AAW18801
ID AAW18801 standard; peptide; 7 AA.
XX
XX AAW18801;
XX
XX 28-JUL-1997 (first entry)
XX
XX RGD peptide for enhancement of LIPOPECTAMINE transfections.
XX
XX transfection; cationic lipid; aggregate; peptide-nucleic acid; PNA;
KW intracellular delivery agent; extracellular targeting; transgenic;
KW animal; gene therapy; viral inhibition; cancer treatment; viral;
KW Spermine-modified; enhancement.
XX
XX Synthetic.
XX
XX WO9640961-A1.
XX
XX 19-DEC-1996.
PD

```

```

XX 04-JUN-1996; 96WO-US008723.
XX 07-JUN-1995; 95US-00477354.
XX (LIFE-) LIFE TECHNOLOGIES INC.
XX Hawley-Nelson P, Lan J, Shih P, Jesse JA, Schifferli KP;
XX WPI; 1997-052346/05.
XX Transfecting eukaryotic cell with peptide-nucleic acid complex - useful
XX in cancer treatment, gene therapy and viral inhibition.
XX Example 5; Page 30; 62pp; English.
XX By adding a cationic lipid, capable of aggregating a peptide-nucleic acid
XX (PNA), to a PNA complex, a composition for transfecting eukaryotic cells
XX is formed. The claimed compositions can be used as intracellular delivery
XX agents and extracellular targeting agents, pref. to transfect eukaryotic
XX cells. The transfected cells which express useful gene prods. can be
XX employed in the prodn. of transgenic animals. The compts. can also be
XX used to introduce nucleic acids into cells in methods of gene therapy and
XX viral inhibition and for introduction of antisense or antigen or related
XX inhibitory nucleic acids into cells, esp. useful in cancer treatment. The
XX present sequence is an RGD peptide. Relatively minor enhancements of
XX transfection compared to the control were observed except for HApep and
XX the combination of VSVG and E5. Spermatize-modified peptides caused even
XX more significant enhancement in the efficiency of cationic lipid
XX transfection. See AAW18796-805
XX Sequence 7 AA;
  Query Match      100.0%; Score 43; DB 2; Length 7;
  Best Local Similarity 100.0%; Pred. No. 1.4e+06;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 GRGDSPC 7
  Db 1 GRGDSPC 7
  RESULT 6
  AAY43230
  ID AAY43230 standard; peptide; 7 AA.
  AC AAY43230;
  XX
  DT 13-JAN-2000 (first entry)
  DE RGD-containing peptide #9.
  XX
  KW Nucleic acid delivery vehicle; bifunctional complex; transgene; CFTK;
  KW cell surface targeting; cell surface molecule binding region; integrin;
  KW cystic fibrosis transmembrane regulator; alpha-antitrypsin;
  KW suicide gene; beta-glucocerebrosidase; cell transfection; cell infection;
  KW RGD peptide.
  OS Synthetic.
  XX
  PN WO9940214-A2.
  PD 12-AUG-1999.
  XX
  PF 08-FEB-1999; 99WO-US002680.
  XX
  PR 09-FEB-1998; 98US-00020483.
  PR 09-FEB-1998; 98US-0135092P.
  PR 06-MCV-1998; 98US-0107471P.
  XX
  PA (GENZ ) GENZYME CORP.
  XX
  PI O'riordan C, Romanczuk H, Wadsworth SC;
  XX
  DR WPI; 1999-610583/52.
  XX
  PT Nucleic acid delivery vehicles useful for transfecting and infecting a
  PT target cell.
  XX
  PS Claim 22; Page 39; 118pp; English.
  XX
  CC This sequence represents a RGD-containing peptide that can be used in a
  CC bifunctional complex used in the nucleic acid delivery vehicle (I) of the
  CC invention. (I) is for transfecting and/or infecting a target cell, and
  CC comprises a transgene and a bifunctional complex (B) that targets the
  CC nucleic acid delivery vehicle to the cell surface. (B) comprises a
  CC delivery vehicle binding portion, a cell surface molecule binding portion
  CC (such as this sequence) and a linker connecting them. The delivery
  CC vehicle can be specifically targeted to the cell via the binding to cell
  CC surface molecules. (I) can be used to target cells, which express
  CC integrins such as, HT-29 colon carcinoma cells, lymphocytes and
  CC monocytes, blood platelets, SMC-9C human lung fibroblast, MG163
  CC osteosarcoma cell line, vascular endothelial cells and melanoma cells.
  CC (I) is useful for delivery of nucleic acids encoding CFTK (cystic
  CC fibrosis transmembrane regulator), alpha-antitrypsin, beta-
  CC glucocerebrosidase and suicide genes. The construct increases the

```

CC efficiency of cellular uptake of (I). The constructs also enable the  
 CC transfection/infection of cells that are normally refractory to  
 CC transfection/infection by targeting cell receptors that are present on  
 CC such cells

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 DB 1 GRGDSPC 7

RESULT 7  
 AAY31133  
 ID AAY31133 standard; peptide; 7 AA.

XX AC AAY31133;  
 XX DT 21-OCT-1999 (first entry);  
 XX DE Non-crosslinked protein particle peptide 182.

XX KW Non-crosslinked protein particle; diagnostic; therapy; monodisperse;  
 KW albumin; haemoglobin; nanometer; micrometer; clearance.  
 XX OS Synthetic.  
 XX PN US945033-A.  
 XX PD 31-AUG-1999.

XX PF 12-NOV-1996; 96US-00747137.  
 XX PR 15-JAN-1991; 91US-00641720.  
 XX PR 13-OCT-1992; 92US-00959560.  
 XX PR 01-JUN-1993; 93US-00069831.  
 XX PR 14-MAR-1994; 94US-00212546.

XX PA (HEMO-) HEMOSPHERE INC.

XX PI Yen RCK;

XX PS WPI; 1999-508153/42.

XX PT Non-crosslinked protein particles for therapeutic and diagnostic use.  
 XX Example 22; Col 125-126; 65pp; English.

XX This invention describes a novel aqueous suspension of monodisperse  
 CC particles on non-crosslinked, non-denatured albumin (50-5000 nm) which is  
 CC stable against dissolving upon dilution with an alcohol-free aqueous  
 CC medium. The method involves (a) forming an aqueous solution containing  
 CC albumin and hemoglobin and (b) treating the aqueous solution with an  
 CC alcohol to cause the solution to become turbid. The particles are useful  
 CC as agents for in vivo administration, either of their own administration  
 CC or as a vehicle for other therapeutic or diagnostic agents. The method  
 CC permits the formation of albumin and hemoglobin particles in the  
 CC nanometer and micrometer size range, in a form closer to their natural  
 CC form than the forms of the prior art. The particles therefore constitute  
 CC a more closely controlled agent for in vivo administration, with greater  
 CC ease of clearance from the body after their period of usefulness.  
 CC AAY30952-Y31-35 represent peptides used in the method of the invention

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 DB 1 GRGDSPC 7

RESULT 8  
 AAY80477  
 ID AAY80477 standard; peptide; 7 AA.

XX AC AAY80477;  
 XX DT 06-JUN-2000 (first entry)  
 XX DE Cell adhesion peptide #12.

XX KW Bone regenerative; osteopathic; osseous tissue; reconstitution;  
 KW scaffold matrix; bone formation promoter; bone resorption inhibitor;  
 KW cell adhesion; osteoblast; osteoclast; bone defect; fracture.

XX OS Synthetic.

XX PN WO200004941-A1.

XX PD 03-FEB-2000.

XX PF 22-JUL-1999; 99WO-US016800.

XX PR 24-JUL-1998; 98US-00122348.

XX PA (PEAR-) PHARMACAL BIOTECHNOLOGIES INC.

XX PI Budny JA;

XX DR WPI; 2000-195084/17.

XX PT System for reconstructing osseous tissue, useful e.g. for treating  
 PT fractures, comprises scaffold containing promoter of bone formation and  
 PT inhibitor of bone resorption.

XX PS Claim 14; Page 31; 44pp; English.

XX The invention relates to a novel system for reconstruction of osseous  
 CC tissue comprising a scaffold carrying a compound (I) that promotes bone  
 CC formation and a component that decreases bone resorption (II). (I)  
 CC induces migration and adhesion of osteoblasts and osteoclasts and (II)  
 CC inhibits proteolysis (specifically by plasmin) of extracellular matrix.  
 CC (I) is preferably selected from: selectin or selectin binding fragments,  
 CC proteins and peptides that facilitate cell adhesion, plasminogen  
 CC activator inhibitors, protease inhibitors and metalloprotease inhibitors.  
 CC The peptides AAY80466-Y80492 are claimed examples of cell adhesion  
 CC peptides used in the system of the invention. The system is used to  
 CC replace, remodel or correct bone defects, e.g. fractures, fissures or  
 CC bone mass loss. Incorporation of (I) into the scaffold results in rapid  
 CC seeding by osteoblasts and the development of an organic matrix, i.e. the  
 CC preformed scaffold replaces the rate-determining step of extracellular  
 CC matrix formation. The scaffold can be designed to have a predetermined  
 CC resorption/degradation rate, and may include regulatory compounds for  
 CC specific cell types

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 3; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 DB 1 GRGDSPC 7

RESULT 9  
 AAB91971  
 ID AAB91971 standard; peptide; 7 AA.



XX AAB91971;  
AC  
XX  
DT 22-JUN-2001 (first entry)  
XX  
XX Fibronectin fragment and fibrin related peptide SEQ ID NO:1147.  
DE  
XX  
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX Homo sapiens.  
OS Synthetic.  
OS  
XX WO200069900-A2.  
XX  
XX 23-NOV-2000.  
XX  
XX 17-MAY-2000; 2000WO-US013576.  
XX  
XX 17-MAY-1999; 99US-C134406P.  
PR 10-SEP-1999; 99US-0153426P.  
PR 15-OCT-1999; 99US-0159783P.  
XX  
XX (CONJ-) CONJUCHEM INC.  
XX  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
PI  
XX WPI; 2001-112059/12.  
DR  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity.  
PT  
XX  
XX Disclosure; Page 570; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (iii) and a  
CC reactive group (ii) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (iv), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity in  
CC vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention  
XX  
XX Sequence 7 AA;  
SQ  
Query Match 100.0%; Score 43; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
DB 1 GRGDSPC 7  
RESULT 10  
AAE28378  
ID AAE28378 standard; peptide; 7 AA.  
XX  
XX AAE28378;  
AC  
XX  
XX 27-DEC-2002 (first entry)  
DT  
XX Fibronectin RGD peptide used in the invention.  
DE

XX Tat region; nucleic acid-binding group; cell transfection system;  
KW gene therapy; cancer; fibronectin; RGD peptide.  
XX  
XX Unidentified.  
OS  
XX US6376248-B1.  
PR  
XX 23-APR-2002.  
PD  
XX 16-MAR-1998; 98US-00039780.  
PF  
XX 14-MAR-1997; 97US-00818200.  
PR  
XX (LIFE-) LIFE TECHNOLOGIES INC.  
PA  
XX Hawley-Nelson P, Ian J, Shih P, Jesse JA, Schifferli KP;  
PI Gebeyehu G, Ciccarone VC, Evans KL;  
PI  
XX WPI; 2002-680647/73.  
DR  
XX New peptide comprising Tat sequence linked to nucleic acid-binding group,  
PT useful, e.g. in gene therapy, for improving cell-transfection efficiency.  
PT  
XX Example 1; Col 57; 108pp; English.  
PS  
XX The invention relates to a peptide comprising Tat sequence linked to  
CC nucleic acid-binding group. Peptides of the invention are used as  
CC components of a cell transfection system particularly for gene therapy  
CC (especially of cancer). The present sequence is fibronectin RGD peptide  
CC used in the invention  
XX  
XX Sequence 7 AA;  
SQ  
Query Match 100.0%; Score 43; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
DB 1 GRGDSPC 7  
RESULT 11  
AAB09833  
ID AAB09833 standard; peptide; 7 AA.  
XX  
XX AAB09833;  
AC  
XX 30-JUL-2002 (first entry)  
DT  
XX Amino acid sequence of peptide #15.  
DE  
XX Osteogenic implant; titanium; bone implant; screw; dentistry.  
KW  
XX Unidentified.  
OS  
XX WO200207792-A2.  
FN  
XX 31-JAN-2002.  
PD  
XX 24-JUL-2001; 2001WO-CH000456.  
PF  
XX 26-JUL-2000; 2000CH-00001481.  
PR  
XX (STRA ) INST STRAUMANN AG.  
PA  
XX Simpson JP, Steinemann SG;  
PI  
XX WPI; 2002-268955/31.  
DR  
XX Titanium osteogenic implant, for use in bone implants, has its  
PT hydroxylated surface coated with 2 or more primary or secondary amino,  
PT

Fri Apr 16 11:04:07 2004

CC in the study, diagnosis, treatment or prevention of diseases and  
 CC conditions which relate to cell adhesion, such as rheumatoid arthritis,  
 CC asthma, allergy, adult respiratory distress syndrome (ARDS), inflammatory  
 CC bowel diseases, ophthalmic inflammatory diseases, autoimmune diseases,  
 CC thrombosis, platelet aggregation conditions, cardiovascular diseases,  
 CC neoplastic diseases, wound healing conditions, and prosthetic  
 CC implantations conditions, reocclusion following thrombolysis, and  
 CC allograft rejection. (Updated on 25-MAR-2003 to correct PA field.)  
 XX Sequence 8 AA;  
 SQ

Query Match 100.0%; Score 43; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 DB 2 GRGDSPC 8

RESULT 13  
 AAR21005  
 ID AAR21005 standard; peptide; 8 AA.  
 XX AC AAR21005;  
 XX DT 25-MAR-2003 (revised)  
 DT 17-OCT-1992 (first entry)  
 XX  
 DE Cyclised integrin receptor antagonists for modulating cell adhesion.  
 XX Integrin; fibronectin; cell adhesion; leukocyte; rgd.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1. .8  
 FT Modified-site 2 /label= NME-Gly  
 FT  
 XX MO9200995-A.  
 PN 23-JAN-1992.  
 XX 09-JUL-1990; 90US-00550330.  
 XX 09-JUL-1990; 90US-00550330.  
 XX (TANA ) TANABE SEIYAKU CO.  
 XX Lobi T, Chiang SL, Cardarelli PM;  
 PT WPI; 1992-056823/07.  
 DR Cyclised peptide receptor antagonists - modulate fibronectin-related and  
 XX leukocyte adhesion to endothelial cells, in treatment of inflammatory, or  
 PT neoplastic diseases.  
 XX Claim 14; Page 102; 128pp; English.  
 XX The peptide is a specifically claimed example of a group of generic  
 CC cyclic peptides having activity as cell adhesion modulators. Some of the  
 CC peptides contain Arg-Gly-Asp sequences. Others contain Arg-Cys- Asp  
 CC sequences, or contain "reversed orientation" (or "retro") peptide  
 CC sequences in which the normal amino-to-carboxy direction of the amino  
 CC acid in the peptide backbone has been reversed. The compounds  
 CC sufficiently mimic extracellular matrix ligands or other cell adhesion  
 CC ligands so as to bind to cell surface receptors, including integrin  
 CC receptors such as the fibronectin, collagen, laminin, LFA-1, MAF-1,  
 CC p150,95, vitronectin and gp11b/IIla receptors. The peptides can be used  
 CC in the study, diagnosis, treatment or prevention of diseases and  
 CC conditions which relate to cell adhesion, such as rheumatoid arthritis,  
 CC asthma, allergy, adult respiratory distress syndrome (ARDS), inflammatory

PT carboxyl, amide, phosphono or OH groups.  
 XX Claim 6; Page 25; 31pp; German.  
 XX The specification describes an osteogenic implant, made of titanium metal  
 CC or a titanium-based alloy. The implant has a roughened surface that, in a  
 CC hydroxylated state, is at least partially covered with a compound  
 CC containing at least 2 groups selected from primary or secondary amino,  
 CC carboxyl, amide, phosphono or OH groups. The implant is used as a bone  
 CC implant, e.g. in the form of a screw for use in dentistry. The present  
 CC sequence represents a peptide that is used in the course of the invention  
 XX

Query Match 100.0%; Score 43; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 DB 1 GRGDSPC 7

RESULT 12  
 AAR21014  
 ID AAR21014 standard; peptide; 8 AA.  
 XX AC AAR21014;  
 XX DT 25-MAR-2003 (revised)  
 DT 17-OCT-1992 (first entry)  
 XX  
 DE Cyclised integrin receptor antagonists for modulating cell adhesion.  
 XX Integrin; fibronectin; cell adhesion; leukocyte; rgd.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1. .8  
 FT Modified-site 1 /note= "Ada-Cys, where Ada = 1-adamantaneacetyl"  
 FT  
 XX MO9200995-A.  
 PN 23-JAN-1992.  
 XX 09-JUL-1990; 90US-00550330.  
 XX 09-JUL-1990; 90US-00550330.  
 XX (TANA ) TANABE SEIYAKU CO.  
 XX Lobi T, Chiang SL, Cardarelli PM;  
 PT WPI; 1992-056823/07.  
 DR Cyclised peptide receptor antagonists - modulate fibronectin-related and  
 XX leukocyte adhesion to endothelial cells, in treatment of inflammatory, or  
 PT neoplastic diseases.  
 XX Claim 14; Page 103; 128pp; English.  
 XX The peptide is a specifically claimed example of a group of generic  
 CC cyclic peptides having activity as cell adhesion modulators. Some of the  
 CC peptides contain Arg-Gly-Asp sequences. Others contain Arg-Cys- Asp  
 CC sequences, or contain "reversed orientation" (or "retro") peptide  
 CC sequences in which the normal amino-to-carboxy direction of the amino  
 CC acid in the peptide backbone has been reversed. The compounds  
 CC sufficiently mimic extracellular matrix ligands or other cell adhesion  
 CC ligands so as to bind to cell surface receptors, including integrin  
 CC receptors such as the fibronectin, collagen, laminin, LFA-1, MAF-1,  
 CC p150,95, vitronectin and gp11b/IIla receptors. The peptides can be used

CC bowel diseases, ophthalmic inflammatory diseases, autoimmune diseases,  
 CC thrombosis, platelet aggregation conditions, cardiovascular diseases,  
 CC neoplastic diseases, wound healing conditions, and prosthetic  
 CC implantations conditions, reocclusion following thrombolysis, and  
 CC allograft rejection. (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 8 AA;  
 Query Match 100.0%; Score 43; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSPC 7  
 DB 2 GRGDSPC 8  
 RESULT 14  
 AAW48503  
 ID AAW48503 standard; peptide; 8 AA.  
 XX AC AAW48503;  
 XX DT 18-AUG-1998 (first entry)  
 XX DE Integrin receptor antagonist peptide 42.  
 XX KW Integrin receptor antagonist; cell adhesion modulator; leukocyte;  
 XX KW extracellular matrix; fibronectin; ARDS; thrombosis; inflammation.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Disulfide-bond 1..8  
 FT Modified-site 1 /note= "3-Mercaptopropionic acid"  
 FT Modified-site 2 /note= "Optionally attached by 1-Adamantaneacetic acid"  
 FT Modified-site 7 /note= "Optionally N-methyl-Glycine"  
 FT Modified-site 8 /note= "Optionally 4-Hydroxyproline"  
 FT Modified-site 8 /note= "Optional C-terminal amide"  
 XX PN US5721210-A.  
 XX PD 24-FEB-1998.  
 XX PF 07-JUN-1995; 95US-00485019.  
 XX PR 09-JUL-1990; 90US-00550330.  
 XX PR 09-JUL-1991; 91WO-US004862.  
 XX PR 04-JUN-1993; 93US-00961889.  
 XX PA (TANA ) TANABE SEIYAKU CO.  
 XX PI Cardarelli PM, Lobl TJ, Chiang S;  
 XX WI; 1998-168442/15.  
 XX PT New cyclic peptide(s) and peptidomimetic compounds - are integrin  
 XX PT receptor antagonists useful in modulating cell adhesion.  
 XX PS Example 9; Col 42; 32pp; English.  
 XX The present sequence represents a synthetic peptide which acts as an  
 XX antagonist to integrin receptors. The invention provides various  
 XX synthetic peptides which act as cell adhesion modulators because they  
 XX mimic extra-cellular matrix ligands or other cell adhesion ligands that  
 XX bind to receptors such as integrin receptors, including fibronectin,  
 XX laminin, LFA-1, MAC-1, p150,95, vitronectin and gp1b/IIb receptors.  
 XX Some of the peptides contain the amino acid sequence Arg-Gly-Asp (RGD).  
 XX Others contain non-RGD sequences, for e.g RGD sequences, and reverse  
 XX orientation forms of amino acid residues. The synthetic peptides are  
 XX useful in modulating cell adhesion, including adhesion related to  
 XX fibronectin, as well as leukocyte adhesion to endothelial cells. They are  
 XX also claimed to be useful in the study, diagnosis, treatment or  
 XX prevention of diseases which relate to cell adhesion, e.g. adult-  
 XX respiratory distress syndrome (ARDS), thrombosis and inflammatory  
 XX conditions

XX SQ Sequence 8 AA;  
 Query Match 100.0%; Score 43; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSPC 7  
 DB 2 GRGDSPC 8  
 RESULT 15  
 AAW48495  
 ID AAW48495 standard; peptide; 8 AA.  
 XX AC AAW48495;  
 XX DT 18-AUG-1998 (first entry)  
 XX DE Integrin receptor antagonist peptide 34.  
 XX KW Integrin receptor antagonist; cell adhesion modulator; leukocyte;  
 XX KW extracellular matrix; fibronectin; ARDS; thrombosis; inflammation.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Disulfide-bond 1..8  
 FT Modified-site 1 /note= "Optionally attached by 1-Adamantaneacetic acid"  
 FT Modified-site 2 /note= "Optionally N-methyl-Glycine"  
 FT Modified-site 7 /note= "Optionally 4-Hydroxyproline"  
 FT Modified-site 8 /note= "Optional C-terminal amide"  
 XX PN US5721210-A.  
 XX PD 24-FEB-1998.  
 XX PF 07-JUN-1995; 95US-00485019.  
 XX PR 09-JUL-1990; 90US-00550330.  
 XX PR 09-JUL-1991; 91WO-US004862.  
 XX PR 04-JUN-1993; 93US-00961889.  
 XX PA (TANA ) TANABE SEIYAKU CO.  
 XX PI Cardarelli PM, Lobl TJ, Chiang S;  
 XX WI; 1998-168442/15.  
 XX PT New cyclic peptide(s) and peptidomimetic compounds - are integrin  
 XX PT receptor antagonists useful in modulating cell adhesion.  
 XX PS Example 9; Col 42; 32pp; English.  
 XX The present sequence represents a synthetic peptide which acts as an  
 XX antagonist to integrin receptors. The invention provides various  
 XX synthetic peptides which act as cell adhesion modulators because they  
 XX mimic extra-cellular matrix ligands or other cell adhesion ligands that  
 XX bind to receptors such as integrin receptors, including fibronectin,  
 XX laminin, LFA-1, MAC-1, p150,95, vitronectin and gp1b/IIb receptors.  
 XX Some of the peptides contain the amino acid sequence Arg-Gly-Asp (RGD).  
 XX Others contain non-RGD sequences, for e.g RGD sequences, and reverse  
 XX orientation forms of amino acid residues. The synthetic peptides are  
 XX useful in modulating cell adhesion, including adhesion related to  
 XX fibronectin, as well as leukocyte adhesion to endothelial cells. They are  
 XX also claimed to be useful in the study, diagnosis, treatment or  
 XX prevention of diseases which relate to cell adhesion, e.g. adult  
 XX respiratory distress syndrome (ARDS), thrombosis and inflammatory  
 XX conditions

Fri Apr 16 11:04:07 2004

CC conditions  
XX Sequence 8 AA;  
SQ  
Query Match 100.0%; Score 43; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GRGDSPC 7  
Db 2 GRGDSPC 8  
Search completed: April 16, 2004, 07:20:56  
Job time : 78.475 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 22.575 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-10

Perfect score: 43

Sequence: 1 GRGDSPC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptcdat2/2/aaa/5A-COMB.pep:\*
- 2: /cgn2\_6/ptcdat2/2/aaa/5B-COMB.pep:\*
- 3: /cgn2\_6/ptcdat2/2/aaa/6A-COMB.pep:\*
- 4: /cgn2\_6/ptcdat2/2/aaa/6B-COMB.pep:\*
- 5: /cgn2\_6/ptcdat2/2/aaa/PCTUS-COMB.pep:\*
- 6: /cgn2\_6/ptcdat2/2/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	100.0	7	1	US-08-178-482-2
2	43	100.0	7	1	US-08-558-130-10
3	43	100.0	7	2	US-08-747-137-182
4	43	100.0	7	3	US-09-426-680-9
5	43	100.0	7	4	US-09-039-780A-8
6	43	100.0	7	5	PCT-US96-08723-10
7	43	100.0	9	1	US-08-445-745-113
8	43	100.0	9	4	US-09-366-991-11
9	43	100.0	9	4	US-08-456-466-91
10	43	100.0	9	4	US-08-445-638-113
11	43	100.0	10	1	US-08-311-835-2
12	43	100.0	10	1	US-08-169-743-6
13	43	100.0	10	1	US-08-445-745-44
14	43	100.0	10	1	US-08-445-745-45
15	43	100.0	10	1	US-08-421-702A-46
16	43	100.0	10	1	US-08-482-106-1
17	43	100.0	10	1	US-08-303-052A-73
18	43	100.0	10	1	US-08-421-696A-46
19	43	100.0	10	1	US-08-482-107B-2
20	43	100.0	10	1	US-08-421-537A-46
21	43	100.0	10	1	US-08-421-538A-46
22	43	100.0	10	2	US-08-421-595A-74
23	43	100.0	10	4	US-09-366-991-1
24	43	100.0	10	4	US-09-366-991-8
25	43	100.0	10	4	US-08-456-466-18
26	43	100.0	10	4	US-08-456-466-19
27	43	100.0	10	4	US-08-445-638-44

Sequence 45, Appli  
Sequence 7, Appli  
Patent No. 5318899  
Sequence 120, App  
Sequence 118, App  
Sequence 119, App  
Sequence 62, Appl  
Sequence 114, App  
Sequence 92, Appl  
Sequence 114, App  
Sequence 50, Appl  
Sequence 102, App  
Sequence 103, App  
Sequence 104, App  
Sequence 105, App  
Sequence 106, App  
Sequence 110, App

ALIGNMENTS

RESULT 1

US-08-178-482-2  
; Sequence 2, Application US/08178482  
; Patent No. 5629294  
; GENERAL INFORMATION:  
; APPLICANT: DIZERGA, GERE S  
; APPLICANT: ROZGERS, KATHLEEN E  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING  
; TITLE OF INVENTION: ADHESION FORMATION  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBBINS, DALGARN, BERLINER & CARSON  
; STREET: 201 NORTH FIGUEROA STREET, FIFTH FLOOR  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90012-2628  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/178,482  
; FILING DATE: 06-JAN-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/789,231  
; FILING DATE: 07-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SPITALS, JOHN P  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-314  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-178-482-2

Query Match 100.0%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7

Db 1 GRGDSPC 7

us-09-991-588b-10.ra1

Fri Apr 16 11:04:07 2004

```

; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,137
; FILING DATE: 12-NOV-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/212,546
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION NUMBER: US 08/069,831
; FILING DATE: 01-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,560
; FILING DATE: 13-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/641,720
; FILING DATE: 15-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 016197-000840US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; INFORMATION FOR SEQ ID NO: 182:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; US-08-747-137-182

Query Match 100.0%; Score 43; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 4
US-09-426-680-9
; Sequence 9, Application US/09426680
; Patent No. 6287857
; GENERAL INFORMATION:
; APPLICANT: Catherine R. O'Riordan
; APPLICANT: Samuel C. Wadsworth
; TITLE OF INVENTION: Nucleic Acid Delivery Vehicles
; FILE REFERENCE: GAO103USB2
; CURRENT APPLICATION NUMBER: US/09/426,680
; CURRENT FILING DATE: 1999-10-25
; EARLIER APPLICATION NUMBER: PCT/US99/02680
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 7
; TYPE: PRT
; ORGANISM: human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; US-09-426-680-9

Query Match 100.0%; Score 43; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 3
US-08-747-137-182
; Sequence 182, Application US/08747137
; Patent No. 5945033
; GENERAL INFORMATION:
; APPLICANT: Yen, Richard C.K.
; TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
; THERAPEUTIC AND DIAGNOSTIC USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco

US-08-658-130-10
; Sequence 10, Application US/08658130
; Patent No. 5736392
; GENERAL INFORMATION:
; APPLICANT: Hawley-Nelson, Pamela
; APPLICANT: Lan, Jiangirg
; APPLICANT: Shih, Po-Jen
; APPLICANT: Jesse, Joel A.
; APPLICANT: Shifferli, Kevin P.
; TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid
; TITLE OF INVENTION: Transfections
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,130
; FILING DATE: 04-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 32-95A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-658-130-10

Query Match 100.0%; Score 43; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
Db 1 GRGDSPC 7

RESULT 3
US-08-747-137-182
; Sequence 182, Application US/08747137
; Patent No. 5945033
; GENERAL INFORMATION:
; APPLICANT: Yen, Richard C.K.
; TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
; THERAPEUTIC AND DIAGNOSTIC USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
```

Fri Apr 16 11:04:07 2004

us-09-991-588b-10.ra1

STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/08723  
FILING DATE: 04-JUN-1996

## CLASSIFICATION:

PRIOR APPLICATION DATA: US 08/477,354  
APPLICATION NUMBER: 07-JUN-1995  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:

NAME: Caruthers, Jennie M.  
REGISTRATION NUMBER: 34,464  
REFERENCE/DOCKET NUMBER: 3295A WO

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089

## INFORMATION FOR SEQ ID NO: 10:

## SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

PCT-US96-08723-10

Query Match 100.0%; Score 43; DB 5; Length 7;

Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels

Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GRGDSPC 7

DB 1 GRGDSPC 7

## RESULT 7

US-08-445-745-113

Sequence 113, Application US/08445745

Patent No. 5672585

## GENERAL INFORMATION:

APPLICANT: Pierschbacher, Michael D.

APPLICANT: Cheng, Soan

APPLICANT: Craig, William S.

APPLICANT: Tschoop, Juerg F.

TITLE OF INVENTION: Methods and Composition for Treating

NUMBER OF SEQUENCES: 168

CORRESPONDENCE ADDRESS:

ADDRESSES: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/445,745

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/171,068

QY 1 GRGDSPC 7

DB 1 GRGDSPC 7

## RESULT 5

US-09-039-780A-8

Sequence 8, Application US/09039780A

Patent No. 6376248

## GENERAL INFORMATION:

APPLICANT: HAMLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERL, KEVIN P.

GEBEYEHU, GUILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,780A

FILING DATE: 16-Mar-1998

## CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95C

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

## INFORMATION FOR SEQ ID NO: 8:

## SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

## SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 100.0%; Score 43; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels

Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GRGDSPC 7

DB 1 GRGDSPC 7

## RESULT 6

PCT-US96-08723-10

Sequence 10, Application PC/TUS9608723

## GENERAL INFORMATION:

APPLICANT: Life Technologies, Inc.

TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

us-09-991-588b-10.ra1

Fri Apr 16 11:04:07 2004

```

OTHER INFORMATION: Xaa=Penicillamine (Pen)
US-09-366-991-11
Query Match 100.0%; Score 43; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 3 GRGDSPC 9

RESULT 9
US-08-456-466-91
; Sequence 91, Application US/08456466
; Patent No. 6395873
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; Lukeman, David S.
; Cheng, Soan
; Craig, William S.
; Tschoopp, Juerg F.
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR TREATING
; THROMBOSIS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSER: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/456,466
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa=Pmp"
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-08-456-466-91
Query Match 100.0%; Score 43; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 2 GRGDSPC 8

RESULT 10
US-08-445-638-113
; Sequence 113, Application US/08445638

```

```

FILING DATE: 20-DEC-1993
APPLICATION NUMBER: US 08/079,441
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,73614
FILING DATE: 14-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,119
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,444
FILING DATE: 06-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9829
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: circular
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note= "Xaa = (Pmp)"
US-08-445-745-113
Query Match 100.0%; Score 43; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 2 GRGDSPC 8

RESULT 8
US-09-366-991-11
; Sequence 11, Application US/09366991
; Patent No. 6353090
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; Ruoslahti, Erkki I.
; TITLE OF INVENTION: Conformationally Stabilized Cell
; Adhesion Peptides
; FILE REFERENCE: P-LA 3637
; CURRENT APPLICATION NUMBER: US/09/366,991
; CURRENT FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 08/459,445
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 08/292,568
; PRIOR FILING DATE: 1994-08-19
; PRIOR APPLICATION NUMBER: US 08/215,012
; PRIOR FILING DATE: 1994-03-21
; PRIOR APPLICATION NUMBER: US 08/124,992
; PRIOR FILING DATE: 1993-09-21
; PRIOR APPLICATION NUMBER: US 08/048,576
; PRIOR FILING DATE: 1993-04-15
; PRIOR APPLICATION NUMBER: US 07/903,797
; PRIOR FILING DATE: 1991-12-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: 2

```



us-09-991-588b-10.ra1

Fri Apr 16 11:04:07 2004

```

; Patent No. 6521594
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; TITLE OF INVENTION: Thrombosis
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,638
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,068
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: US 08/079,441
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,73614
; FILING DATE: 14-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/506,444
; FILING DATE: 06-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9000
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa = (Pmp)."
; US-08-445-638-113

Query Match 100.0%; Score 43; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 2 GRGDSPC 8

RESULT 11
US-08-311-835-2
; Sequence 2, Application US/08311835
; Patent No. 5493007
; GENERAL INFORMATION:
; APPLICANT: Burnier, John P.
; APPLICANT: Gadek, Thomas,

; Patent No. 6521594
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; TITLE OF INVENTION: Thrombosis
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,638
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,068
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: US 08/079,441
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,73614
; FILING DATE: 14-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/506,444
; FILING DATE: 06-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9000
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa = (Pmp)."
; US-08-445-638-113

Query Match 100.0%; Score 43; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 2 GRGDSPC 8

RESULT 11
US-08-311-835-2
; Sequence 2, Application US/08311835
; Patent No. 5493007
; GENERAL INFORMATION:
; APPLICANT: Burnier, John P.
; APPLICANT: Gadek, Thomas,

; Patent No. 6521594
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Cheng, Soan
; APPLICANT: Craig, William S.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: Methods and Composition for Treating
; TITLE OF INVENTION: Thrombosis
; NUMBER OF SEQUENCES: 168
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,638
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,068
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: US 08/079,441
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,73614
; FILING DATE: 14-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/506,444
; FILING DATE: 06-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9000
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa = (Pmp)."
; US-08-445-638-113

Query Match 100.0%; Score 43; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
DB 3 GRGDSPC 9

RESULT 12
US-08-163-743-6
; Sequence 6, Application US/08169743
; Patent No. 5547936
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: GEHLEN, KURT R.
; TITLE OF INVENTION: INHIBITION OF CELL MIGRATION WITH
; TITLE OF INVENTION: SYNTHETIC PEPTIDES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/169,743
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/773,106
FILING DATE: 08-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, THERESA A.
REGISTRATION NUMBER: 32,547
REFERENCE/DOCKET NUMBER: P319102
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /note= "X=Pen"
US-08-169-743-6
Query Match 100.0%; Score 43; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 1 GRGDSPC 7
DD 3 GRGDSPC 5

RESULT 13
US-08-445-745-44
Sequence 44, Application US/08445745
Patent No. 5672585
GENERAL INFORMATION:
APPLICANT: Pierschbacher, Michael D.
APPLICANT: Cheng, Soan
APPLICANT: Craig, William S.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: Methods and Composition for Treating
TITLE OF INVENTION: Thrombosis
NUMBER OF SEQUENCES: 168
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,745
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,068
FILING DATE: 22-DEC-1993
APPLICATION NUMBER: US 08/079,441
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,73614
FILING DATE: 14-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,119

```

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 2  
OTHER INFORMATION: /note= "Xaa is Pen"  
US-08-421-702A-46  
Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
Db 3 GRGDSPC 9  
Search completed: April 16, 2004, 07:31:40  
Job time : 22.575 secs

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9829  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1  
OTHER INFORMATION: /note= "Xaa = R or K"  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 2  
OTHER INFORMATION: /note= "Xaa = Per"  
US-08-445-745-45

Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
Db 3 GRGDSPC 9

RESULT 15  
US-08-421-702A-46  
Sequence 46, Application US/08421702A  
Patent No. 5759396  
GENERAL INFORMATION:  
APPLICANT: Cheng, Soan  
APPLICANT: Ingram, Ronald  
APPLICANT: Mullen, Daniel  
APPLICANT: Tschoopp, Juerg  
TITLE OF INVENTION: Peptides Useful for Altering Alpha-  
TITLE OF INVENTION: v Beta-3-Mediated Blinding  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/421,702A  
FILING DATE: 12-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/303,052  
FILING DATE: 08-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 1480  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:25:03 ; Search time 26.9 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588B-1  
Perfect score: 25  
Sequence: 1 RGDC 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	4	9 US-09-765-614B-1	Sequence 1, Appli
2	26	100.0	4	9 US-09-925-715-4	Sequence 4, Appli
3	26	100.0	4	11 US-09-991-588B-1	Sequence 1, Appli
4	26	100.0	4	15 US-10-179-743A-1	Sequence 37, Appli
5	26	100.0	5	9 US-09-364-597A-37	Sequence 9, Appli
6	26	100.0	5	12 US-09-912-609-9	Sequence 2, Appli
7	26	100.0	5	14 US-10-304-160-2	Sequence 7, Appli
8	26	100.0	6	9 US-09-364-597A-7	Sequence 10, Appli
9	26	100.0	6	12 US-09-912-609-10	Sequence 13, Appli
10	26	100.0	6	12 US-09-912-609-13	Sequence 5, Appli
11	26	100.0	7	9 US-09-823-444-5	Sequence 12, Appli
12	26	100.0	7	9 US-09-840-277-12	Sequence 50, Appli
13	26	100.0	7	9 US-09-840-277-50	Sequence 148, App
14	26	100.0	7	10 US-09-791-524-148	Sequence 322, App
15	26	100.0	7	12 US-10-609-217-322	

16	26	100.0	7	12 US-10-632-338-322	Sequence 322, App
17	26	100.0	7	12 US-10-415-324-29	Sequence 29, Appli
18	26	100.0	7	12 US-10-651-723-322	Sequence 322, App
19	26	100.0	7	14 US-10-131-543-29	Sequence 29, Appli
20	26	100.0	7	14 US-10-131-546-29	Sequence 29, Appli
21	26	100.0	7	14 US-10-131-346-29	Sequence 29, Appli
22	26	100.0	8	9 US-09-946-833-9	Sequence 9, Appli
23	26	100.0	8	12 US-10-013-C09-4	Sequence 4, Appli
24	26	100.0	8	14 US-10-046-801-23	Sequence 23, Appli
25	26	100.0	9	9 US-09-765-086-1	Sequence 1, Appli
26	26	100.0	9	9 US-09-845-160-5	Sequence 5, Appli
27	26	100.0	9	9 US-09-245-603A-16	Sequence 16, Appli
28	26	100.0	9	9 US-09-364-597A-15	Sequence 15, Appli
29	26	100.0	9	9 US-09-364-597A-16	Sequence 16, Appli
30	26	100.0	9	9 US-09-364-597A-17	Sequence 17, Appli
31	26	100.0	9	9 US-09-364-597A-18	Sequence 18, Appli
32	26	100.0	9	9 US-09-364-597A-33	Sequence 33, Appli
33	26	100.0	9	9 US-09-734-628-1	Sequence 1, Appli
34	26	100.0	9	9 US-09-971-798-5	Sequence 5, Appli
35	26	100.0	9	9 US-09-969-192-3	Sequence 3, Appli
36	26	100.0	9	9 US-09-969-192-4	Sequence 4, Appli
37	26	100.0	9	9 US-09-840-277-14	Sequence 14, Appli
38	26	100.0	9	9 US-09-840-277-22	Sequence 22, Appli
39	26	100.0	9	9 US-09-840-277-38	Sequence 38, Appli
40	26	100.0	9	9 US-09-840-277-62	Sequence 62, Appli
41	26	100.0	9	9 US-09-840-277-63	Sequence 63, Appli
42	26	100.0	9	10 US-09-801-485-2	Sequence 2, Appli
43	26	100.0	9	10 US-09-791-524-124	Sequence 124, App
44	26	100.0	9	12 US-10-609-217-449	Sequence 449, App
45	26	100.0	9	12 US-10-609-217-450	Sequence 450, App

## ALIGNMENTS

RESULT 1  
US-09-765-614B-1  
; Sequence 1, Application US/09765614B  
; Patent No. US20020102215A1  
; GENERAL INFORMATION:  
; APPLICANT: Nycomed Imaging AS  
; TITLE OF INVENTION: Improvements in or relating to  
; TITLE OF INVENTION: diagnostic/therapeutic  
; TITLE OF INVENTION: agents  
; FILE REFERENCE: REF/KIavness/054  
; CURRENT APPLICATION NUMBER: US/09/765,614B  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial  
; OTHER INFORMATION: Sequence:RGDC-Mal-PEG3400-DSPE  
US-09-765-614B-1

Query Match 100.0%; Score 26; DB 9; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06; 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
Db 1 RGDC 4

RESULT 2  
US-09-925-715-4  
; Sequence 4, Application US/09925715  
; Patent No. US20020102217A1  
; GENERAL INFORMATION:  
; APPLICANT: Nycomed Imaging AS

;; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic  
;; TITLE OF INVENTION: agents  
;; FILE REFERENCE: REF/Klaveness/206  
;; CURRENT APPLICATION NUMBER: US/09/925,715  
;; CURRENT FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 27  
;; SOFTWARE: Patentin Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 4  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence:Synthetic RGDC  
;; OTHER INFORMATION: sequence  
US-09-925-715-4

Query Match 100.0%; Score 26; DB 5; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 3

US-09-991-588B-1  
;; Sequence 1, Application US/09991588B  
;; Publication No. US20030219429A1  
;; GENERAL INFORMATION:

;; APPLICANT: Budny, John A.  
;; TITLE OF INVENTION: Compositionand Method for Bone Regeneration  
;; FILE REFERENCE: 1008-120.US  
;; CURRENT APPLICATION NUMBER: US/09/991,588B  
;; CURRENT FILING DATE: 2001-11-21  
;; PRIOR APPLICATION NUMBER: US 09/122,348  
;; PRIOR FILING DATE: 1998-07-24  
;; NUMBER OF SEQ ID NOS: 23  
;; SEQ ID NO 1  
;; LENGTH: 4  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:

;; OTHER INFORMATION: Purchased commercially or sequence is synthesized  
US-09-991-588B-1

Query Match 100.0%; Score 26; DB 11; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 4

US-10-179-743A-1  
;; Sequence 1, Application US/10179743A  
;; Publication No. US20040001959A1  
;; GENERAL INFORMATION:

;; APPLICANT: Schwartz, Jeffrey  
;; APPLICANT: Danahy, Michael P  
;; APPLICANT: Gawalt, Ellen S.  
;; APPLICANT: Schwarzbauer, Jean E.  
;; APPLICANT: Midwood, Kim S.  
;; APPLICANT: Avalltroni, Michael J.  
;; APPLICANT: Carolus, Michael D.  
;; TITLE OF INVENTION: ENHANCED BONDING LAYERS ON TITANIUM MATERIALS

;; FILE REFERENCE: P22,571-1 USA  
;; CURRENT APPLICATION NUMBER: US/10/179,743A  
;; CURRENT FILING DATE: 2002-06-24  
;; PRIOR APPLICATION NUMBER: US 60/389,574  
;; PRIOR FILING DATE: 2002-06-18

;; PRIOR APPLICATION NUMBER: US 60/300,144  
;; PRIOR FILING DATE: 2001-06-22  
;; PRIOR APPLICATION NUMBER: US 09/668,980  
;; PRIOR FILING DATE: 2000-09-22  
;; PRIOR APPLICATION NUMBER: US 60/155,398  
;; PRIOR FILING DATE: 1999-09-22  
;; PRIOR APPLICATION NUMBER: US 08/794,833  
;; PRIOR FILING DATE: 1997-02-04  
;; PRIOR APPLICATION NUMBER: US 60/035,940  
;; PRIOR FILING DATE: 1997-01-13  
;; PRIOR APPLICATION NUMBER: US 60/028,949  
;; PRIOR FILING DATE: 1996-10-17  
;; NUMBER OF SEQ ID NOS: 1  
;; SOFTWARE: Patentin version 3.1  
;; SEQ ID NO 1  
;; LENGTH: 4  
;; TYPE: PRT  
;; ORGANISM: Artificial  
;; FEATURE:  
;; OTHER INFORMATION: The three amino acid sequence RGD which occurs on Fibronectine c  
;; OTHER INFORMATION: 11 adhesion protein is modified with cystine to provide RGDC  
US-10-179-743A-1

Query Match 100.0%; Score 26; DB 15; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
|||  
Db 1 RGDC 4

## RESULT 5

US-09-364-597A-37  
;; Sequence 37, Application US/09364597A  
;; Patent No. US20020103130A1  
;; GENERAL INFORMATION:

;; APPLICANT: Ruoslahti, Erkki  
;; APPLICANT: Koivunen, Erkki  
;; TITLE OF INVENTION: No. US20020103130A1e1 Integrin-Binding Peptides  
;; NUMBER OF SEQUENCES: 46  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell & Flores LLP  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92122

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/364,597A  
;; FILING DATE: 30-JUL-1999  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/158,001  
;; FILING DATE: 24-NOV-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/286,861  
;; FILING DATE: 04-AUG-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-LA 3419  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (858) 535-9001  
;; TELEFAX: (858) 535-8949  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: circular
US-09-364-597A-37

Query Match      100.0%; Score 26; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
   ||||
DB 2 RGDC 5

RESULT 6
US-09-912-609-9
; Sequence 9, Application US/09912609
; Publication No. US20020041898A1
; GENERAL INFORMATION:
; APPLICANT: UNGER, EVAN C.
; APPLICANT: MATSUNAGA, TERRY ONICHI
; APPLICANT: RAMASWAMI, VARADARAJAN
; APPLICANT: ROMANOWSKI, MAREK J.
; TITLE OF INVENTION: NOVEL TARGETED DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 5030-0001.24
; CURRENT APPLICATION NUMBER: US/09/912,609
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 09/703,474
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/478,124
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-912-609-9

Query Match      100.0%; Score 25; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
   ||||
DB 2 RGDC 5

RESULT 7
US-10-304-160-2
; Sequence 2, Application US/10304160
; Publication No. US20030099619A1
; GENERAL INFORMATION:
; APPLICANT: WICKHAM, THOMAS J
; APPLICANT: KOVESDI, IMRE
; APPLICANT: ROELVINK, PETROS W
; APPLICANT: EINFELD, DAVID
; APPLICANT: BROUGH, DOUGLAS E
; APPLICANT: LIZONOVA, ALENA
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR TARGETING AN ADENOVIRAL VECTOR
; FILE REFERENCE: 220148
; CURRENT APPLICATION NUMBER: US/10/304,160
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: PCT/US01/17391
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 09/631,191
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: US 60/208451
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-304-160-2

Query Match      100.0%; Score 26; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
   ||||
DB 2 RGDC 5

RESULT 8
US-09-364-597A-7
; Sequence 7, Application US/09364597A
; Patent No. US20020103130A1
; GENERAL INFORMATION:
; APPLICANT: Ruuslahti, Erkki
; APPLICANT: Koivunen, Erkki
; TITLE OF INVENTION: No. US20020103130A1e1 Integrin-Binding Peptides
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/364,597A
; FILING DATE: 30-JUL-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/158,001
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,861
; FILING DATE: 04-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 3419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (858) 535-9001
; TELEFAX: (858) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-09-364-597A-7

Query Match      100.0%; Score 26; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
   ||||
DB 2 RGDC 5

RESULT 9
US-09-912-609-10
; Sequence 10, Application US/09912609
; Publication No. US20020041898A1
```

```
; GENERAL INFORMATION:
; APPLICANT: UNGER, EVAN C.
; APPLICANT: MATSUNAGA, TERRY ONICHI
; APPLICANT: RAMASWAMI, VARADARAJAN
; APPLICANT: ROMANOWSKI, MAREK J.
; TITLE OF INVENTION: NOVEL TARGETED DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 5030-0001.24
; CURRENT APPLICATION NUMBER: US/09/912,609
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 09/703,474
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/478,124
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-912-609-10

Query Match          100.0%; Score 26; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 2 RGDC 5

RESULT 10
US-09-912-609-13
; Sequence 13, Application US/09912609
; Publication No. US20020041898A1
; GENERAL INFORMATION:
; APPLICANT: UNGER, EVAN C.
; APPLICANT: MATSUNAGA, TERRY ONICHI
; APPLICANT: RAMASWAMI, VARADARAJAN
; APPLICANT: ROMANOWSKI, MAREK J.
; TITLE OF INVENTION: NOVEL TARGETED DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 5030-0001.24
; CURRENT APPLICATION NUMBER: US/09/912,609
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 09/703,474
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/478,124
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-912-609-13

Query Match          100.0%; Score 26; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 2 RGDC 5
```

```
; Patent No. US20020009753A1
; GENERAL INFORMATION:
; APPLICANT: Bednar, Bohumil
; APPLICANT: Boliag, Daniel M.
; APPLICANT: Gould, Robert J.
; APPLICANT: Merck & Co., Inc.
; TITLE OF INVENTION: ANTICOAGULANT TEST
; FILE REFERENCE: 19910
; CURRENT APPLICATION NUMBER: US/09/823,444
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This sequence is a synthetically prepared peptide.
US-09-823-444-5

Query Match          100.0%; Score 26; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 4 RGDC 7

RESULT 12
US-09-840-277-12
; Sequence 12, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RGD, NGR derivative peptide
; NAME/KEY: misc feature
; LOCATION: (2 and)..(3)
; OTHER INFORMATION: Xaa is any amino acid
US-09-840-277-12

Query Match          100.0%; Score 26; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 4 RGDC 7

RESULT 13
US-09-840-277-50
; Sequence 50, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
```

```
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-698A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Integrin antagonist peptide
; NAME/KEY: misc:feature
; LOCATION: (2)..(3)
; OTHER INFORMATION: Xaa is any amino acid residue
US-09-840-277-50

Query Match      100.0%; Score 26; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      |||||
        4 RGDC 7

RESULT 14
US-09-791-524-148
; Sequence 148, Application US/09791524
; Publication No. US20030143209A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharmaceuticals Products Inc.
; TITLE OF INVENTION: Targeted Adenovirus Vectors For Delivery Of Heterologous Genes
; FILE REFERENCE: A3319A
; CURRENT APPLICATION NUMBER: US/09/791,524
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/09828
; PRIOR FILING DATE: 1998-08-27
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 148
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Adenovirus
US-09-791-524-148

Query Match      100.0%; Score 26; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      |||||
        3 RGDC 6

RESULT 15
US-10-609-217-322
; Sequence 322, Application US/10609217
; Publication No. US20040044188A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
```

```
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/609,217
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 322
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: INHIBITION OF PLATELET AGGREGATION
; NAME/KEY: misc:feature
; LOCATION: (2)..(3)
; OTHER INFORMATION: Xaa = any amino acid
US-10-609-217-322

Query Match      100.0%; Score 26; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDC 4
Db      |||||
        4 RGDC 7

Search completed: April 16, 2004, 08:03:15
Job time : 26.9 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 8.4 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-1

Perfect score: 26

Sequence: 1 RGDC 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	69	2 T35541	Hypothetical prote
2	26	100.0	82	2 I42502	CI4L protein - vac
3	26	100.0	100	2 C72598	hypothetical prote
4	26	100.0	103	1 K4RB	Ig kappa-B4 chain
5	26	100.0	103	2 T49368	hypothetical prote
6	26	100.0	104	1 K3RB	Ig kappa-B9 chain
7	26	100.0	104	2 A69037	conserved hypot het
8	26	100.0	106	1 K4RBS	Ig kappa-2 chain C
9	26	100.0	106	2 G20907	Ig kappa-B4 chain
10	26	100.0	109	2 C72588	hypothetical prote
11	26	100.0	130	2 S57896	polyketide synthas
12	26	100.0	131	2 A80899	conserved hypot het
13	26	100.0	140	2 S47353	ribosomal protein
14	26	100.0	140	2 A75141	hypothetical prote
15	26	100.0	141	2 G71099	hypothetical prote
16	26	100.0	142	2 A13246	conserved hypot het
17	26	100.0	145	2 D69383	conserved hypot het
18	26	100.0	147	2 I54245	laminin M - rat (f
19	26	100.0	160	2 T10838	probable resistanc
20	26	100.0	161	2 S65777	pathogenesis-relat
21	26	100.0	161	2 B85354	PR-1-like protein
22	26	100.0	161	2 B90392	conserved hypot het
23	26	100.0	162	2 T08154	pathogenesis-relat
24	26	100.0	162	2 H83792	hypothetical prote
25	26	100.0	169	2 H72470	hypothetical prote
26	26	100.0	172	1 A28539	C-phycoyanin 2 be
27	26	100.0	172	2 S00713	phycoyanin 1 beta
28	26	100.0	172	2 S75318	phycoyanin beta c
29	26	100.0	173	2 S05711	phycoyanin 3 beta

## ALIGNMENTS

### RESULT 1

T35541  
hypothetical protein SC6G10.32 - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C;Accession: T35541  
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: Z21581  
A;Accession: T35541  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-69 <SEE>  
A;Cross-references: EMBL:AL049497; PIDN:CAB39886.1; GSPDB:GN00070; SCOEDB:SC6G10.32  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SC6G10.32

Query Match 100.0%; Score 26; DB 2; Length 69;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4  
Db 22 RGDC 25

### RESULT 2

I42502  
CI4L protein - vaccinia virus (strain Copenhagen)  
C;Species: vaccinia virus  
A;Note: host Homo sapiens (man)  
C;Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 15-Sep-2003  
C;Accession: I42502  
R;Johnson, G.P.  
submitted to GenBank, June 1990  
A;Reference number: A33172  
A;Accession: I42502  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-82 <JUH>  
C;Superfamily: uncharacterized conserved protein

Query Match 100.0%; Score 26; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4  
Db 39 RGDC 42

## RESULT 3

C72598  
 hypothetical protein APE1251 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C:Accession: C72598  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hailkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; X DNA Res. 6, 83-102, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: C72598  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-100 <XAW>  
 A:Cross-references: DDBJ:AP000061; NID:G5104821; PIDN:BA80241.1; PID:dl044027; PID:G5104821  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE1251  
 C:Superfamily: Aeropyrum pernix hypothetical protein APE1251

Query Match 100.0%; Score 26; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4

DB 41 RGDC 44

## RESULT 4

K4RB  
 Ig kappa-B4 chain C region - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 31-Jan-1981 #sequence\_revision 15-Oct-1982 #text\_change 16-Aug-1996  
 C:Accession: A93971; A93891; A92176; A92122  
 R:Emorine, L.; Dreher, K.; Kindt, T.J.; Max, E.E.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 5709-5713, 1983  
 A:Title: Rabbit immunoglobulin kappa genes: structure of a germline b4 allotype J-C locus  
 A:Reference number: A93971; MUID:83300036; PMID:6412231  
 A:Accession: A93971  
 A:Molecule type: DNA  
 A:Residues: 1-103 <EMO>  
 A:Note: the sequence was determined from the germline gene  
 R:Heidmann, O.; Aulfray, C.; Cazenave, P.A.; Rougeon, F.  
 Proc. Natl. Acad. Sci. U.S.A. 78, 5802-5806, 1981  
 A:Title: Nucleotide sequence of constant and 3' untranslated regions of a kappa immunoglobulin  
 A:Reference number: A93891; MUID:82060334; PMID:6795636

A:Accession: A93891  
 A:Molecule type: mRNA  
 A:Residues: 1-103 <HEI>  
 R:Chen, K.C.S.; Kindt, T.J.; Krause, R.M.  
 J. Biol. Chem. 250, 3289-3296, 1975  
 A:Title: Primary structure of the L chain from a rabbit homogeneous antibody to streptococcus  
 A:Reference number: A92176; MUID:75133568; PMID:1091650  
 A:Accession: A92176  
 A:Molecule type: protein  
 A:Residues: 1-57, D', 59-103 <CHE>  
 A:Note: this chain was obtained from antibody to the specific carbohydrate of group C streptococcus  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:19-87/Domain: immunoglobulin homology <IMV>

Query Match 100.0%; Score 26; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4

DB 100 RGDC 103

## RESULT 5

T49368  
 hypothetical protein B1D1.220 [imported] - Neurospora crassa  
 C:Species: Neurospora crassa  
 C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
 C:Accession: T49368  
 R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022  
 A:Accession: T49368  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <SCH>  
 A:Cross-references: EMBL:AL355927; GSPDB:GN00116; NCSP:B1D1.220  
 A:Experimental source: BAC clone B1D1; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B1D1.220  
 A:Map position: 6  
 C:Superfamily: Neurospora crassa hypothetical protein B1D1.220

Query Match 100.0%; Score 26; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4

DB 90 RGDC 93

## RESULT 6

K3RB  
 Ig kappa-B9 chain C region - rabbit (tentative sequence)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 31-Mar-2000  
 C:Accession: A02120  
 R:Parasuram, V.; Goodfellow, R.; Rodkey, S.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 73, 1293-1296, 1976  
 A:Title: Immunoglobulin allotypes of rabbit kappa chains: polymorphism of a control mechanism  
 A:Reference number: A02120; MUID:76176480; PMID:817288

A:Accession: A02120  
 A:Molecule type: protein  
 A:Residues: 1-104 <FAR>  
 A:Note: this chain was obtained from pooled immunoglobulin of homozygous B9 rabbits  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:20-08/Domain: immunoglobulin homology <IMV>

Query Match 100.0%; Score 26; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4

DB 101 RGDC 104

## RESULT 7

A69037  
 conserved hypothetical protein MTH1278 - Methanobacterium thermoautotrophicum (strain D)  
 C:Species: Methanobacterium thermoautotrophicum  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: A69037  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Liu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: A69037  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-104 <VTE>  
A:Cross-references: GB:AE000893; GB:AE000666; NID:G2622375; PID:AB85760.1; PID:G262238  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MPH1278  
A:Start codon: GTG

Query Match 100.0%; Score 26; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
DB 48 RGDC 51

RESULT 8  
IG kappa-2 chain C region (K-bas) - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 24-Sep-1999  
C:Accession: A02121; A29115  
R:Heidmann, O.; Rougeon, F.  
EMBO J. 2, 437-441, 1983  
A:Title: Multiplicity of constant kappa light chain genes in the rabbit genome: a b4b4 H  
A:Reference number: A02121  
A:Accession: A02121  
A:Molecule type: DNA  
A:Residues: 1-106 <HEI>  
A:Experimental source: strain Basilea, b4 haplotype (C-kappa-1 locus)  
R:Marlame, B.; Akimenko, M.A.; Rougeon, F.  
Nucleic Acids Res. 15, 6171-6179, 1987  
A:Title: Interallelic and intergenic conversion events could induce differential evolution  
A:Reference number: A29115; MUID:87315877; PMID:3114714  
A:Accession: A29115  
A:Molecule type: DNA  
A:Residues: 1-106 <MAR>  
A:Cross-references: GB:X05800; NID:G2176; PID:CAA29246.1; PID:G984675  
A:Experimental source: strain Basilea, b6 and b9 haplotypes (C-kappa-1 locus)  
C:Comment: Although there are two loci encoding kappa C region, the expression of this C  
ession of the C-kappa-1 locus is suppressed, and this form is expressed as a minor compo  
C:Genetics:  
A:Gene: C-kappa-2; K-bas  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hair disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: heterotetramer  
F:20-89/Domain: immunoglobulin homology <IMM>  
F:27-87/Disulfide bonds: #status predicted

Query Match 100.0%; Score 26; DB 1; Length 106;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
DB 103 RGDC 106

RESULT 9  
IG kappa-B4 chain C region (K-bas) - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 21-Jan-2000  
C:Accession: G20907  
R:Emorine, L.; Max, E.E.  
Nucleic Acids Res. 11, 8877-8890, 1983  
A:Title: Structural analysis of a rabbit immunoglobulin kappa2 J-C locus reveals multipl  
A:Reference number: A20907; MUID:84169523; PMID:6324107  
A:Accession: G20907  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-106 <EMO>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:20-89/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 26; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
DB 103 RGDC 106

RESULT 10  
C72598  
hypothetical protein APE1176 - Aeropyrum pernix (strain KI)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: C72598  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: C72598  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-109 <KAW>  
A:Cross-references: DDBJ:AP000061; NID:G5104821; PID:BAA80161.1; PID:dl043947; PID:G510  
A:Experimental source: strain KI  
C:Genetics:  
A:Gene: APE1176

Query Match 100.0%; Score 26; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
DB 80 RGDC 83

RESULT 11  
S57896  
polyketide synthase A - Aspergillus parasiticus (fragments)  
N:Contains: acyl carrier protein; acyltransferase; ketoacylsynthase  
C:Species: Aspergillus parasiticus  
C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 17-Mar-1999  
C:Accession: S57896  
R:Chang, P.K.; Cary, J.W.; Yu, J.; Bhatnagar, D.; Cleveland, T.E.  
Mol. Gen. Genet. 248, 270-277, 1995  
A:Title: The Aspergillus parasiticus polyketide synthase gene pksA, a homolog of Aspergi  
A:Reference number: S57896; MUID:96004693; PMID:7565588  
A:Accession: S57896  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-130 <CHA>  
A:Cross-references: EMBL:Z47198  
A:Experimental source: ATCC 56775  
C:Genetics:  
A:Gene: pksA

Query Match 100.0%; Score 26; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
DB 37 RGDC 40

RESULT 12

## AE0899

Conserved hypothetical protein yxaN [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C;Species: Salmonella enterica subsp. enterica serovar Typhimurium  
 A;Note: This species has also been called Salmonella typhimurium  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: A30899  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Karpman, A.; Davis, P.; Davies, R.A.; Dowd, J.; White, N.; Farrar, M.; Mouton, S.; O'Garra, P.  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A;Reference number: A30899; MUID:21534947; PMID:11677608  
 A;Accession: AE0899  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-131 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD07787.1; PID:G16504335; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: yxaN  
 C;Superfamily: hypothetical protein H11656

Query Match 100.0%; Score 26; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4  
 Db 7 RGDC 10

## RESULT 13

S47353

Ribosomal protein L18a, cytosolic - human

C;Species: Homo sapiens (man)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999

C;Accession: S47353

R;Zenz, K.I.

submitted to the EMBL Data Library, August 1994

A;Reference number: S47353

A;Accession: S47353

A;Molecule type: mRNA

A;Residues: 1-140 &lt;ZEN&gt;

A;Cross-references: EMBL:X80822; NID:g527577; PIDN:CAA56788.1; PID:g527578

C;Superfamily: rat ribosomal protein L18a

C;Keywords: cytosol; protein biosynthesis; ribosome

Query Match 100.0%; Score 26; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4  
 Db 23 RGDC 26

## RESULT 14

A75141

hypothetical protein PAB0623 - Pyrococcus abyssi (strain Orsay)

C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C;Accession: A75141

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure

A;Reference number: A75001

A;Accession: A75141

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-140 &lt;KAW&gt;

A;Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CA849842.1; PID:g545835

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB0623

C;Superfamily: Methanococcus jannaschii conserved hypothetical protein M20206

Query Match 100.0%; Score 26; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4  
 Db 33 RGDC 36

## RESULT 15

G71099

hypothetical protein PH1059 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000

C;Accession: G71099

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki,

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: G71099

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-141 &lt;KAW&gt;

A;Cross-references: GB:APC00004; NID:g3236131; PIDN:BAA30157.1; PID:g3257474

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1059

C;Superfamily: Methanococcus jannaschii conserved hypothetical protein M20206

Query Match 100.0%; Score 26; DB 2; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RGDC 4  
 Db 34 RGDC 37

Search completed: April 16, 2004, 37:29:16  
 Job time : 10.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time 7 Seconds  
(without alignments)  
29.754 Million cell updates/sec

Title: US-09-991-588B-1  
Perfect score: 26  
Sequence: 1 RGDC 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 3  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	82	1 VAC14 VACCC	P21045 vaccinia vi
2	26	100.0	103	1 KAC4 RABIT	P01840 oryctolagus
3	26	100.0	104	1 KAC9 RABIT	P01838 oryctolagus
4	26	100.0	106	1 KACB RABIT	P01839 oryctolagus
5	26	100.0	131	1 YRAN SALTI	Q82318 salmonella
6	26	100.0	131	1 YRAN SALTY	P50203 ambystoma m
7	26	100.0	143	1 HXA9 AMBME	Q98924 gallus gall
8	26	100.0	169	1 HXA9 CHICK	P07119 fremyella d
9	26	100.0	172	1 PHB1 PREDI	Q52446 pseudanabae
10	26	100.0	172	1 PHB1 PSEA9	P08039 fremyella d
11	26	100.0	172	1 PHB2 PREDI	Q52451 pseudanabae
12	26	100.0	172	1 PHB2 PSEA9	Q54714 synchocyst
13	26	100.0	173	1 PHB3 PREDI	P14877 fremyella d
14	26	100.0	173	1 PHB3 PSEA9	Q94866 mus musculu
15	26	100.0	190	1 CTAB MOUSE	P17365 vaccinia vi
16	26	100.0	190	1 VAC13 VACCV	Q96kr6 homo sapien
17	26	100.0	192	1 CTAB HUMAN	O27724 methanobact
18	26	100.0	217	1 YG85 METHA	Q57995 methanococc
19	26	100.0	229	1 YG75 METJA	P73664 synchocyst
20	26	100.0	237	1 SFSA SYNY3	O87322 rhizobium m
21	26	100.0	239	1 SFSA RHIME	Q8ysk0 anabaena sp
22	26	100.0	241	1 SFSA ANASP	Q8di93 synchococc
23	26	100.0	241	1 SFSA SNEBL	P25576 saccharomyc
24	26	100.0	258	1 YCE7 YEAST	P09631 mus musculu
25	26	100.0	271	1 HXA9 MOUSE	P11269 homo sapien
26	26	100.0	272	1 HXA9 HUMAN	Q00946 african swi
27	26	100.0	273	1 VPRT ASFEB7	Q65228 african swi
28	26	100.0	273	1 VPRT ASFEM2	Q91583 pseudomonas
29	26	100.0	312	1 YR58 PSEAE	O28115 archaeoglob
30	26	100.0	313	1 YL67 ARCFU	Q9jhb0 mus musculu
31	26	100.0	338	1 ANX9 MOUSE	P92172 caenorhabdi
32	26	100.0	350	1 DAF7 CAEEL	Q81bl1 arabidopsis
33	26	100.0	363	1 KCO1 ARATH	

34	26	100.0	366	1 SPSY MOUSE	P97355 mus musculu
35	26	100.0	366	1 VUL HSVGU	Q01349 human herpe
36	26	100.0	378	1 RAPA BACSU	Q00828 bacillus su
37	26	100.0	380	1 SPB7 MOUSE	Q04695 mus musculu
38	26	100.0	381	1 SIR2 LEIMA	Q25337 leishmania
39	26	100.0	387	1 YBJF PASMU	Q9cph1 pasteurella
40	26	100.0	390	1 DAF1 MOUSE	Q61475 mus musculu
41	26	100.0	396	1 EFIA GIALA	Q08046 giardia lam
42	26	100.0	398	1 NUCM CAFRO	Q9ta77 cafeteria r
43	26	100.0	400	1 NUCM PROWI	Q37619 prototecta
44	26	100.0	405	1 ARGJ PSESM	Q87wz4 p arginine
45	26	100.0	416	1 PXN1 XENLA	P49263 xenopus lae

## ALIGNMENTS

RESULT 1					
VC14 VACCC					
ID VAC14 VACCC	STANDARD;	PRT;	82 AA.		
AC P21045;					
DT 01-FEB-1991 (Rel. 17, Created)					
DT 01-FEB-1991 (Rel. 17, Last sequence update)					
DT 16-OCT-2001 (Rel. 40, Last annotation update)					
DE Protein Cl4.					
GN Cl4L.					
OS Vaccinia virus (strain Copenhagen).					
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;					
OC Orthopoxvirus.					
OX NCBI_TaxID=10249;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=91021027; PubMed=2219722;					
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,					
RA Paoletti E.;					
RT "The complete DNA sequence of vaccinia virus.";					
RL Virology 179:247-266(1990).					
RN [2]					
RP COMPLETE GENOME.					
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,					
RA Paoletti E.;					
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";					
RL Virology 179:517-563(1990).					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
CC EMBL; M35027; AAA47982.1; -					
DR PIR; I42502; I42502.					
DR PIRSF; PIRSF003692; VAC CL4L; 1.					
SQ SEQUENCE 82 AA; 9304 MW; 83521F7153796C5C CRC64;					
Query Match 100.0%; Score 26; DB 1; Length 82;					
Best Local Similarity 100.0%; Pred. No. 25;					
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 1 RGDC 4					
DB 39 RGDC 42					
RESULT 2					
KAC4 RABIT					
ID KAC4 RABIT	STANDARD;	PRT;	103 AA.		
AC P01840;					
DT 21-JUL-1986 (Rel. 01, Created)					
DT 21-JUL-1986 (Rel. 01, Last sequence update)					
DT 15-MAR-2004 (Rel. 43, Last annotation update)					

DE Ig kappa-B4 chain C region.  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83300036; PubMed=6412231;  
 RA Emorine L., Dreher K.L., Kindt T.J., Max E.E.;  
 RT "Rabbit immunoglobulin kappa genes: structure of a germline b4  
 RT allotype J-C locus and evidence for several b4-related sequences in  
 RT the rabbit genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713 (1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82060134; PubMed=6795636;  
 RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;  
 RT "Nucleotide sequence of constant and 3' untranslated regions of a  
 RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806 (1981).  
 RN [3]  
 RP SEQUENCE.  
 RX MEDLINE=75133568; PubMed=1091650;  
 RA Chen K.C.S., Kindt T.J., Krause R.M.;  
 RT "Primary structure of the L chain from a rabbit homogeneous antibody  
 RT to streptococcal carbohydrate. II. Sequence determination of peptides  
 RT from tryptic and peptic digests.";  
 RL J. Biol. Chem. 250:3289-3296 (1975).  
 CC -!- MISCELLANEOUS: This chain was obtained from antibody to the  
 CC specific carbohydrate of group C Streptococci and was isolated  
 CC from the serum of a single rabbit.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; X00231; CAA25051.1; -;  
 DR PIR; A33971; K9RB.  
 DR HSSP; P01942; 7FAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00407; IGc1; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 5 95 IG-LIKE.  
 FT DISULFID 26 85  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT CONFLICT 58 58 N -> D (IN REF. 3).  
 SQ SEQUENCE 103 AA; 11043 MW; 5FC5ACC8B60E69EB CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 100 RGDC 103  
 RESULT 3  
 ID XAC9 RABIT STANDARD; PRT; 104 AA.  
 AC P01942;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DE Ig kappa-B4 chain C region.  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83300036; PubMed=6412231;  
 RA Emorine L., Dreher K.L., Kindt T.J., Max E.E.;  
 RT "Rabbit immunoglobulin kappa genes: structure of a germline b4  
 RT allotype J-C locus and evidence for several b4-related sequences in  
 RT the rabbit genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713 (1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82060134; PubMed=6795636;  
 RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;  
 RT "Nucleotide sequence of constant and 3' untranslated regions of a  
 RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806 (1981).  
 RN [3]  
 RP SEQUENCE.  
 RX MEDLINE=75133568; PubMed=1091650;  
 RA Chen K.C.S., Kindt T.J., Krause R.M.;  
 RT "Primary structure of the L chain from a rabbit homogeneous antibody  
 RT to streptococcal carbohydrate. II. Sequence determination of peptides  
 RT from tryptic and peptic digests.";  
 RL J. Biol. Chem. 250:3289-3296 (1975).  
 CC -!- MISCELLANEOUS: This chain was obtained from antibody to the  
 CC specific carbohydrate of group C Streptococci and was isolated  
 CC from the serum of a single rabbit.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; X00231; CAA25051.1; -;  
 DR PIR; A33971; K9RB.  
 DR HSSP; P01942; 7FAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00407; IGc1; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 5 95 IG-LIKE.  
 FT DISULFID 26 85  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT CONFLICT 58 58 N -> D (IN REF. 3).  
 SQ SEQUENCE 103 AA; 11043 MW; 5FC5ACC8B60E69EB CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 100 RGDC 103

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ig kappa-B4 chain C region.  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76176480; PubMed=817288;  
 RA Farnsworth V., Goodflesh R., Rodkey S., Hood L.;  
 RT "Immunoglobulin allotypes of rabbit kappa chains: polymorphism of a  
 RT control mechanism regulating closely linked duplicated genes?";  
 RL Proc. Natl. Acad. Sci. U.S.A. 73:1293-1296 (1976).  
 CC -!- MISCELLANEOUS: This chain was obtained from pooled immunoglobulin  
 CC of homozygous b9 rabbits.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC PIR; A02120; K9RB.  
 DR HSSP; P01942; 7FAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00407; IGc1; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 6 100 IG-LIKE.  
 FT DISULFID 27 86  
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
 SQ SEQUENCE 104 AA; 11347 MW; 427B1668B0EC8D98 CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 101 RGDC 104  
 RESULT 4  
 ID XACB RABIT STANDARD; PRT; 106 AA.  
 AC P01939;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DE Ig kappa-B4 chain C region.  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83300036; PubMed=6412231;  
 RA Emorine L., Dreher K.L., Kindt T.J., Max E.E.;  
 RT "Rabbit immunoglobulin kappa genes: structure of a germline b4  
 RT allotype J-C locus and evidence for several b4-related sequences in  
 RT the rabbit genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713 (1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82060134; PubMed=6795636;  
 RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;  
 RT "Nucleotide sequence of constant and 3' untranslated regions of a  
 RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806 (1981).  
 RN [3]  
 RP SEQUENCE.  
 RX MEDLINE=75133568; PubMed=1091650;  
 RA Chen K.C.S., Kindt T.J., Krause R.M.;  
 RT "Primary structure of the L chain from a rabbit homogeneous antibody  
 RT to streptococcal carbohydrate. II. Sequence determination of peptides  
 RT from tryptic and peptic digests.";  
 RL J. Biol. Chem. 250:3289-3296 (1975).  
 CC -!- MISCELLANEOUS: This chain was obtained from antibody to the  
 CC specific carbohydrate of group C Streptococci and was isolated  
 CC from the serum of a single rabbit.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; X00231; CAA25051.1; -;  
 DR PIR; A33971; K9RB.  
 DR HSSP; P01942; 7FAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00407; IGc1; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 5 95 IG-LIKE.  
 FT DISULFID 26 85  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT CONFLICT 58 58 N -> D (IN REF. 3).  
 SQ SEQUENCE 103 AA; 11043 MW; 5FC5ACC8B60E69EB CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDC 4  
 DB 100 RGDC 103

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; V01241; CAA24558.1;
DR EMBL; V00885; -; NCT_ANNOTATED_CDS.
DR PIR; A02121; K4RBBS.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_YHC.
DR Pfam; PF02021; JPF0102; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; FALSE NEG.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 1
FT DOMAIN 6 99 IG-LIKE.
FT DISULFID 27 87
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 106 AA; 11279 MW; AF9B928DDA853849 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
DB 103 RGDC 106

RESULT 5
YRAN SALTY STANDARD; PRT; 131 AA.
AC Q8ZJ18;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical UPF0102 protein yran.
GN YRAN OR STV3448 OR T3185.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=60;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.E., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 703931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- SIMILARITY: Belongs to the UPF0102 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A016844; AAO70723.1; -
DR HAMAP; MF 00048; -; 1.
DR InterPro; IPR003509; UPF0102.
DR Pfam; PF02021; JPF0102; 1.
DR TIGRFAMs; TIGR00252; TIGR00252; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA; 14860 MW; C89B1A87AE02B463 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
DB 7 RGDC 10

RESULT 6
YRAN SALTY STANDARD; PRT; 131 AA.
AC Q8ZJ18;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical UPF0102 protein yran.
GN YRAN OR STM3265.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- SIMILARITY: Belongs to the UPF0102 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A500850; AAL22137.1; -
DR StyGene; SG????; yran.
DR HAMAP; MF 00048; -; 1.
DR InterPro; IPR003509; UPF0102.
DR Pfam; PF02021; UPF0102; 1.
DR TIGRFAMs; TIGR00252; TIGR00252; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA; 14903 MW; E6E2FD19CF94A662 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
DB 7 RGDC 10

```

```

RESULT 7
HXA9 AMBME STANDARD; PRT; 143 AA.
AC P50209;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A9 (Fragment).
GN HXA9.
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Regenerating limb blastema;
RX MEDLINE=95324399; PubMed=7600989;
RA Gardner D.M., Blumberg B., Komine Y., Bryant S.V.;
RT "Regulation of HoxA expression in developing and regenerating axolotl
limbs.";
RL Development 121:1731-1741(1995).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
a developmental regulatory system that provides cells with
specific positional identities on the anterior-posterior axis (By
similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Abd-B homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U20941; AAA86505.1; -.
CC HSSP; P02833; 9ANT.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR006711; Hox9 act.
CC InterPro; IPR000047; HTH_lamb-repressr.
CC Pfam; PF00046; homeobox; 1.
CC Pfam; PF04617; Hox9 act; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON TER 1
FT DNA BIND 78 137 HOMEBOX.
SQ SEQUENCE 143 AA; 16592 MW; 83ED182CBE573773 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
DB 16 RGDC 19

RESULT 8
HXA9 CHICK STANDARD; PRT; 169 AA.
AC Q98924;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Homeobox protein Hox-A9 (Fragment).
GN HXA9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97081281; PubMed=8922326;
RA Gaunt S.J., Strachan L.;
RT "temporal colinearity in expression of anterior Hox genes in
developing chick embryos.";
RL Dev. Dyn. 207:270-280(1996).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
a developmental regulatory system that provides cells with
specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Abd-B homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X97750; CAA66331.1; -.
CC HSSP; P02833; 9ANT.
CC TRANSFAC; TC1708; -.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR006711; Hox9 act.
CC InterPro; IPR000047; HTH_lamb-repressr.
CC Pfam; PF00046; homeobox; 1.
CC Pfam; PF04617; Hox9 act; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON TER 1
FT DNA BIND 103 162 HOMEBOX.
SQ SEQUENCE 169 AA; 19668 MW; 9D74877543924B87 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 169;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
DB 38 RGDC 41

RESULT 9
PHB1 FREDI STANDARD; PRT; 172 AA.
AC P07119;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-phycocyanin-1 beta chain.
GN CPCB1.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
OX NCBI_TaxID=1197;
RN [1]
RP SEQUENCE FROM N.A.
RA Mazel D., Houmard J., Tandeau de Marsac N.;

```



RT "A multigene family in Calothrix sp. PCC 7601 encodes phycocyanin, the major component of the cyanobacterial light harvesting antenna.";  
 RL Mol. Genet. 211:296-304(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8872492; PubMed=3127591;  
 RA Conley P.B., Lemaux P.G., Grossman A.;  
 RT "Molecular characterization and evolution of sequences encoding light-harvesting components in the chromatically adapting cyanobacterium Fremyella diplosiphon.";  
 RL J. Mol. Biol. 199:447-465(1988).  
 RN [3]  
 RP SEQUENCE OF 158-172 FROM N.A.  
 RX MEDLINE=86233345; PubMed=3086870;  
 RA Conley P.B., Lemaux P.G., Lomax T.L., Grossman A.R.;  
 RT "Genes encoding major light-harvesting polypeptides are clustered on the genome of the cyanobacterium Fremyella diplosiphon.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3924-3928(1986).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS).  
 RX MEDLINE=91132654; PubMed=1899708;  
 RA Duerling M., Schmidt G.B., Huber R.;  
 RT "Isolation, crystallization, crystal structure analysis and refinement of constitutive C-phycocyanin from the chromatocally adapting cyanobacterium Fremyella diplosiphon at 1.66-A resolution.";  
 RL J. Mol. Biol. 217:577-592(1991).  
 CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein from the phycobiliprotein complex.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.  
 CC -!- INDUCTION: PHYCOCYANIN-1 IS EXPRESSED AT SIMILAR LEVELS IN GREEN AND RED LIGHT (CONSTITUTIVE PHYCOCYANIN).  
 CC -!- PTM: Contains two covalently linked bilin chromophores.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X06084; CAA29464.1; -;  
 DR EMBL; X07013; CAA30064.1; -;  
 DR EMBL; M13218; AAA24878.1; -;  
 DR PIR; S00713; S00713.  
 DR PDB; 1CPC; 15-JAN-93.  
 DR InterPro; IPR001659; Phycobilisome.  
 DR InterPro; IPR006247; Phycocyanin\_b.  
 DR Pfam; PF00502; Phycobilisome; 1.  
 DR ProDom; PD000340; Phycobilisome; 1.  
 DR TIGRfams; TIGR01339; phycoc beta; 1.  
 DR Phycobilisome; Electron transport; Photosynthesis; Bile pigment; Methylation; Multigene family; 3D-structure.  
 FT MOD\_RES 72 82 METHYLATION (BY SIMILARITY).  
 FT BINDING 82 82 PHYCOCYANOBILIN CHROMOPHORE.  
 FT BINDING 153 153 PHYCOCYANOBILIN CHROMOPHORE.  
 FT CONFLICT 4 4 A -> T (IN REF. 2).  
 FT CONFLICT 63 63 A -> H (IN REF. 2).  
 FT CONFLICT 67 67 I -> N (IN REF. 2).  
 FT CONFLICT 76 76 S -> T (IN REF. 2).  
 FT HELIX 4 13  
 FT TURN 14 16  
 FT TURN 21 32  
 FT TURN 33 33  
 FT TURN 34 46  
 FT HELIX 37 47  
 FT TURN 47 47  
 FT HELIX 48 62  
 FT TURN 64 66  
 FT TURN 67 67  
 FT TURN 69 70  
 FT TURN 72 73  
 FT TURN 76 99  
 FT HELIX 100 100

FT TURN 102 102  
 FT HELIX 103 108  
 FT TURN 109 112  
 FT HELIX 113 120  
 FT TURN 121 121  
 FT HELIX 124 142  
 FT TURN 143 143  
 FT TURN 146 147  
 FT HELIX 154 172  
 SQ SEQUENCE 172 AA; 17920 MW; 5C0E56CS21F0DE1F CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RGDC 4  
 Db 150 RGDC 153  
 RESULT 10  
 PHB1\_PSEA9 STANDARD; PRT; 172 AA.  
 ID PHB1\_PSEA9  
 AC Q52446;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE C-phycocyanin-1 beta chain.  
 OS Pseudanabaena sp. (strain PCC 7409).  
 GN Bacteria; Cyanobacteria; Oscillatoriales;  
 OC Pseudanabaena/Liimothrix group; Pseudanabaena.  
 OX NCBI\_TaxID=29415;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dubbs J.M., Bryant D.A.;  
 RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein from the phycobiliprotein complex.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.  
 CC -!- PTM: Contains two covalently linked bilin chromophores.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M99426; AAA26038.1; -;  
 DR HSP; P07119; ICPC.  
 DR InterPro; IPR001659; Phycobilisome.  
 DR InterPro; IPR006247; Phycocyanin\_b.  
 DR Pfam; PF00502; Phycobilisome; 1.  
 DR ProDom; PD000340; Phycobilisome; 1.  
 DR TIGRfams; TIGR01339; phycoc beta; 1.  
 DR Phycobilisome; Electron transport; Photosynthesis; Bile pigment; Methylation; Multigene family.  
 FT MOD\_RES 72 82 METHYLATION (BY SIMILARITY).  
 FT BINDING 82 82 PHYCOCYANOBILIN CHROMOPHORE.  
 FT BINDING 153 153 PHYCOCYANOBILIN CHROMOPHORE.  
 FT SEQUENCE 172 AA; 17881 MW; 867ADC19E3D85B8F CRC64;  
 Query Match 100.0%; Score 26; DB 1; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RGDC 4  
 Db 150 RGDC 153

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE C-phycoerythrin-2 beta chain.
DE C-PCB2.
OC Pseudanabaena sp. (strain PCC 7409).
OC Bacteria; Cyanobacteria; Oscillatoriales;
OC Pseudanabaena/Limnothrix group; Pseudanabaena.
NCBI_TaxID=29415;
RN [1]
RN SEQUENCE FROM N.A.
RP Dubbs J.M., Bryant D.A.;
RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein
CC from the phycoobiliprotein complex.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- PTM: Contains two covalently linked bilin chromophores.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to licenses@isb-sib.ch).
CC -----
EMBL: M99427; AAA26043.1; -.
DR HSSP: P07119; 1CPC.
DR InterPro: IPR001659; Phycobilisome.
DR InterPro: IPR006247; Phycocyanin b.
DR Pfam: PF00552; Phycobilisome; 1.
DR ProDom: PD003340; Phycobilisome; 1.
DR TIGRFAMs: TIGR01339; phycocy beta; 1.
DR Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Methylation; Multi-gene family.
MOE_RES 72 72 METHYLATION (BY SIMILARITY).
FT BINDING 82 82 PHYCOCYANIN-BILIN CHROMOPHORE.
FT BINDING 153 153 PHYCOCYANOBILIN CHROMOPHORE.
SQ SEQUENCE 172 AA; 18113 MW; C1929DCA5C3E7A74 CRC64;
Query Match 100.0%; Score 26; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDC 4
Db 150 RGDC 153
RESULT 13
PHCB SYN3 STANDARD; PRT; 172 AA.
ID ID PHCB SYN3
AC Q54714; P73205;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 42, Last annotation update)
DE C-phycoerythrin beta chain.
DE C-PCB OR SLI1577.
GN PCPB OR SLI1577.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
NCBI_TaxID=1148;
RN [1]
RN SEQUENCE FROM N.A.
RP Plank T., Anderson L.K.;
RL "Cloning and sequence analysis of the genes encoding CpcB and CpcA
RL from Synechocystis sp. PCC 6803.";
RL (In) Plant Gene Register PGR95-087.
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,
RA Hosoguchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RN [2]

```

```

RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136 (1996).
RN [3]
RX MEDLINE=93384903; PubMed=2506452;
RA Mazel D., Marliere P.;
RT "Adaptive eradication of methionine and cysteine from cyanobacterial
RT light-harvesting proteins.";
RL Nature 341:245-248 (1989).
CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein
CC from the phycobiliprotein complex.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- INDUCTION: PHYCOCYANIN-3 IS EXPRESSED IN RED LIGHT UNDER
CC CONDITIONS OF SULFUR DEPRIVATION.
CC -!- PTM: Contains two covalently linked bilin chromophores.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06083; CAA29459.1; -
CC DR HSSP; P07119; 1CPC.
CC DR InterPro; IPR001659; Phycobilisome.
CC DR InterPro; IPR006247; Phycocyanin_b.
CC DR Pfam; PF00502; Phycobilisome; 1.
CC DR ProDom; PD000340; Phycobilisome; 1.
CC DR TIGRFAMs; TIGR01339; phycocy beta; 1.
CC DR Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
CC KW Methylation; Multigene family.
CC XW MOD_RES 73 73 METHYLATION
CC FT BINDING 83 83 PHYCOCYANOBILIN CHROMOPHORE.
CC FT BINDING 154 154 PHYCOCYANOBILIN CHROMOPHORE.
CC SQ SEQUENCE 173 AA; 18236 MW; 2C564183D2016B8 CRC64;
CC
CC Query Match 100.0%; Score 26; DB 1; Length 173;
CC Best Local Similarity 100.0%; Pred. No. 52;
CC Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;
CC
CC QY 1 RGDC 4
CC DB 151 RGDC 154
CC
CC RESULT 15
CC CTAS MOUSE
CC ID CTAS MOUSE STANDARD; PRT; 190 AA.
CC AC Q3DBH6; Q91WA2;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Protein C20orf108 homolog.
CC GN C20ORF108.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=C57BL/6J; TISSUE=Small intestine;
CC EX MEDLINE=22354683; PubMed=12466851;
CC RA Okazaki Y., Furuno M., Yasukawa T., Adachi J., Bono H., Kondo S.,
CC RA Yagi K., Tomaru Y., Hasegawa Y., Nogi A., Schonbach C., Gojohori T.,
CC RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
CC RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
CC RA Blake J.A., Bradi D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
CC RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
CC RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
CC RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
CC RA Kana A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
CC RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
CC RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
CC RA Nagashima T., Nemata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

```

RA Petrovsky K., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilmink L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland, and Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein V.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guraratne P.H.,  
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman C.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood C., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra V.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AK008190; BAB2521.1; -;  
 DR EMBL; AK076087; BAC36173.1; -;  
 DR EMBL; BC016210; AAH16210.1; -;  
 DR EMBL; BC057642; AAH57642.1; -;  
 DR MGD; MGI:1914267; 2010011120Rik.  
 KW Transmembrane.  
 FT TRANSMEM 21 41 POTENTIAL.  
 FT TRANSMEM 97 117 POTENTIAL.  
 FT TRANSMEM 148 168 POTENTIAL.  
 FT CONFLICT 73 73 G -> S (IN REF. 1).  
 SQ SEQUENCE 190 AA; 20314 MW; 61C6FEB68322B236 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGJC 4  
 |||||  
 DB 55 RGJC 58

Search completed: April 16, 2004, 07:22:09  
 Job time : 9 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 31.6 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-1

Perfect score: 26  
Sequence: 1 RGDC 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- SPTREMBL.25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvrius.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	26	100.0	42	10 Q9AWK9	Q9awk9 oryza sativ
2	26	100.0	55	11 Q9CS36	Q9cs36 mus musculus
3	26	100.0	62	12 Q91UJ0	Q91uj0 aconitum la
4	26	100.0	64	10 Q8SQ09	Q8sq09 oryza sativ
5	26	100.0	64	10 Q84ZN4	Q84zn4 oryza sativ
6	26	100.0	68	12 Q69588	Q69588 human herpe
7	26	100.0	69	16 Q9X817	Q9x817 streptomyce
8	26	100.0	78	16 Q8XJH8	Q8xjh8 rallostonia s
9	26	100.0	79	7 Q9TNY7	Q9tny7 encorhynchu
10	26	100.0	80	6 Q9MZG2	Q9mzg2 callithrix
11	26	100.0	80	7 F79646	F79646 encorhynchu
12	26	100.0	80	7 Q9XRJ7	Q9xrd7 salmo salar
13	26	100.0	80	7 Q9XRJ3	Q9xrd3 salmo salar
14	26	100.0	80	7 Q9XRJ6	Q9xrd6 salmo salar
15	26	100.0	81	2 Q7WXT6	Q7wxt6 alcaligenes
16	26	100.0	82	12 Q9Q3N6	Q9q3n6 hepatitis c

17	26	100.0	83	16 Q8FC30	Q8fc30 escherichia
18	26	100.0	88	12 Q88801	Q88801 eastern equ
19	26	100.0	93	2 Q50307	Q50307 chroococcus
20	26	100.0	93	2 Q50308	Q50308 chroococcus
21	26	100.0	93	2 Q54534	Q54534 chroococcus
22	26	100.0	94	10 Q9L112	Q9l112 oryza sativ
23	26	100.0	98	16 Q89TR4	Q89tr4 bradyrhizob
24	26	100.0	99	10 Q9ANL6	Q9ani6 bradyrhizob
25	26	100.0	99	10 Q9FP01	Q9fp01 oryza sativ
26	26	100.0	100	16 Q89H78	Q89h78 bradyrhizob
27	26	100.0	100	17 Q9YCK6	Q9yck6 aeropyrum p
28	26	100.0	102	15 Q9TBL6	Q9tbl6 human immun
29	26	100.0	104	17 Q27339	Q27339 methanobact
30	26	100.0	105	11 Q9D3E1	Q9d3e1 mus musculu
31	26	100.0	109	6 Q9BGZ9	Q9bgz9 macaca fasc
32	26	100.0	109	13 Q7T217	Q7t217 oreochromis
33	26	100.0	109	17 Q9YCT6	Q9yct6 aeropyrum p
34	26	100.0	111	11 Q8BZ88	Q8bz88 mus musculu
35	26	100.0	113	16 Q8VWD4	Q8vwd4 streptomyce
36	26	100.0	113	16 Q9ANP8	Q9ang8 bradyrhizob
37	26	100.0	114	11 Q9RL51	Q9rl51 cavia porce
38	26	100.0	115	2 Q51817	Q51817 lactobaci
39	26	100.0	115	5 Q9VAK8	Q9vak8 grosophila
40	26	100.0	118	12 Q92517	Q92517 beet virus
41	26	100.0	118	12 Q9YPH2	Q9ypb2 broad bean
42	26	100.0	119	12 Q9IY53	Q9iy53 potato mop-
43	26	100.0	119	12 Q85292	Q85292 potato mop-
44	26	100.0	119	12 Q80QA2	Q80qa2 potato mop-
45	26	100.0	122	4 Q8NAQ9	Q8naq9 homo sapien

## ALIGNMENTS

### RESULT 1

ID	Q9AWK9	PRELIMINARY;	PRT;	42 AA.
AC	Q9AWK9			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	P0436D06.4 protein.			
GN	P0436D06.4.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Nipponbare;			
RA	Sasaki T.; Matsumoto T.; Yamamoto K.;			
RT	*Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC			
RT	clone:P0436D06.4.			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP003054; BAB33268.1; -			
DR	Gramene; Q9AWK9; -			
SQ	SEQUENCE 42 AA; 4412 MW; 6C5269D2E152A050 CRC64;			

Query Match 100.0%; Score 26; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4  
Db 20 RGDC 23

### RESULT 2

ID	Q9CS36	PRELIMINARY;	PRT;	55 AA.
AC	Q9CS36			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fanconi anemia, complementation group A (Fragment).  
 GN FANCA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085663; PubMed=1217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RC "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK019196; BAB31595.1;  
 DR MGD; MGI:1341823; FANCA.  
 FT NON TER 1  
 SQ SEQUENCE 55 AA; 6300 MW; 8E78F4C66AAB3350D CRC64;

Query Match 100.0%; Score 26; DB 11; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 DB 25 RGDC 28

RESULT 3  
 Q91UJ0 PRELIMINARY; PRT; 62 AA.  
 ID Q91UJ0;  
 AC Q91UJ0;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ORF4 protein.  
 OS Aconitum latent virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.  
 OX NCBI\_TaxID=101764;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=D;  
 RA Fuji S., Yamamoto H., Inoue M., Yamashita K., Fukui Y., Furuya H.,  
 RA Naito H.;  
 RT "Complete nucleotide sequence of the genomic RNA of Aconitum latent  
 RT carlavirus isolated from Delphinium sp."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB051848; BAB56117.1;  
 DR InterPro; IPR003411; Coat 7kDa.  
 DR Pfam; PF02495; 7kD coat; I.  
 SQ SEQUENCE 62 AA; 6916 MW; 82FE4E4B6C0D831 CRC64;

Query Match 100.0%; Score 26; DB 12; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 DB 24 RGDC 27

RESULT 4  
 Q8S0U9 PRELIMINARY; PRT; 64 AA.  
 ID Q8S0U9;  
 AC Q8S0U9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE OJ1014\_G12.26 protein (B1144D11.5 protein).  
 GN OJ1014\_G12.26 OR B1144D11.5;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC  
 RT clone:OJ1014\_G12.26";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:BL144D11.1";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003372; BAB9091.1;  
 DR EMBL; AP003443; BAC00658.1;  
 DR Gramene; Q8S0U9;  
 SQ SEQUENCE 64 AA; 6914 MW; 23FD10E79B4CED73 CRC64;

Query Match 100.0%; Score 26; DB 10; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 DB 58 RGDC 61

RESULT 5  
 Q84ZM4 PRELIMINARY; PRT; 64 AA.  
 ID Q84ZM4;  
 AC Q84ZM4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE OJ1559\_F09.22 protein.  
 GN OJ1559\_F09.22  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OJ1559\_F09.22";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003837; BAC57284.1;  
 SQ SEQUENCE 64 AA; 6988 MW; ACDA106B668A26FB CRC64;

Query Match 100.0%; Score 26; DB 13; Length 64;

```

KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 69 AA; 7240 MW; 5033381C5D2C40D CRC64;

Query Match 100.0%; Score 26; DB 16; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 51 RGDC 54

RESULT 6
Q69588 PRELIMINARY; PRT; 68 AA.
AC Q69588;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DE 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SHL1 protein (Fragment).
GN SHL1.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=95146989; PubMed=7844567;
RA Gompels U.A., Macaulay H.A.;
RT "Characterisation of human telomeric repeat sequences from human
herpesvirus-6 and relationship to replication.";
RL J. Gen. Virol. 76:451-458 (1995).
DR EMBL; X79798; CAA56195.1; -.
FT NON_FER 1
SQ SEQUENCE 68 AA; 7297 MW; F8634C6DC423CAA9 CRC64;

Query Match 100.0%; Score 26; DB 12; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 48 RGDC 51

RESULT 7
Q9X817 PRELIMINARY; PRT; 69 AA.
AC Q9X817;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SC02159.
GN SC02159 OR SC6G10.32.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronan A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL919111; CAB39886.1; -.
PR; T35541; T35541.

```

---

```

KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 69 AA; 7240 MW; 5033381C5D2C40D CRC64;

Query Match 100.0%; Score 26; DB 16; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 22 RGDC 25

RESULT 8
Q8XU88 PRELIMINARY; PRT; 78 AA.
AC Q8XU88;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC3210.
GN RSC3210 OR RS02426.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
Chandler M., Choisme N., Moisan A., Claudel-Renard C., Cunnac S., Demange N.,
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AL646074; CADL6998.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 8537 MW; F8974525BE1CB627 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4
Db 67 RGDC 70

RESULT 9
Q9TNY7 PRELIMINARY; PRT; 79 AA.
AC Q9TNY7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MHC class I alpha 2 antigen (Fragment).
OS Oncorhynchus tshawytscha (Chinook salmon) (King salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=74940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ont-JA*3; TISSUE=fin;
RA Miller K.M., Withler R.E.;
RT "The salmonid class I MHC: limited diversity in a primitive teleost.";
RL Immunol. Rev. 0:0-0 (1999).
DR EMBL; AF104591; AAD26087.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR001039; MHC_I.

```

DR Pfam: PF00129; MHC\_I; 1.  
 DR PRINTS; PR01638; MHCCLASSI.  
 DR PRODOM; PD000050; MHC\_I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 79  
 SQ SEQUENCE 79 AA; 8844 MW; 95FCB8EE0CD3FF9B CRC64;

Query Match 100.0%; Score 26; DB 7; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 ||||  
 Db 7 RGDC 10

## RESULT 10

C9MZG0 PRELIMINARY; PRT; 80 AA.  
 AC Q9MZG0;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Endoepine-like protein (Fragment).  
 OS Callithrix jacchus (Common marmoset).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrich.  
 OX NCBI\_TaxID=9483;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20480701; PubMed=11024254;  
 RA Ivell R., Pusch W., Balvers M., Valentin M., Walther N., Weinbauer G.;  
 RT "Progressive inactivation of the haploid expressed gene for the sperm-  
 specific endoepine-like peptide (ELP) through primate evolution.";  
 RL Gene 255:335-345(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bathgate R.A., Ivell R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF229805; AAF79123.1; -.  
 DR HSSP; P07107; IACA.  
 DR GO; GO:0000062; F:acyl-CoA binding; IEA.  
 DR InterPro; IPR000582; Ac\_coa\_bind\_prot.  
 DR Pfam; PF00887; ACBP; 1.  
 DR PRINTS; PR00689; AC0ABINDINGP.  
 DR PRODOM; PD351532; Ac\_coa\_bind\_prot; 1.  
 DR PROSITE; PS00980; ACBP; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 80 AA; 8928 MW; DD79DED6DF48E595 CRC64;

Query Match 100.0%; Score 26; DB 6; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 ||||  
 Db 30 RGDC 33

## RESULT 11

P79646 PRELIMINARY; PRT; 80 AA.  
 AC P79646;  
 DT 01-MAY-1997 (TREMBLrel. 33, Created)  
 DT 01-MAY-1997 (TREMBLrel. 33, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE MHC class I alpha-2 (Fragment).  
 OS Oncomychnus tshawytscha (Chinook salmon) (King salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=74940;  
 RN [1]

RP SEQUENCE OF 3-28 FROM N.A.  
 RC STRAIN=A2-Onts-HN-3; TISSUE=Liver;  
 RX MEDLINE=98009322; PubMed=9348703;  
 RA Miller K.M., Withler R.E., Beacham T.D.;  
 RT "Molecular evolution at Mhc genes in two populations of chinook salmon  
 Mol. Ecol. 6:937-954(1997).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=A2-Onts-HN-3; TISSUE=Liver;  
 RA Miller K.M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U80296; AAC78420.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR001039; MHC\_I.  
 DR Pfam; PF00129; MHC\_I; 1.  
 DR PRINTS; PR01638; MHCCLASSI.  
 DR PRODOM; PD000050; MHC\_I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 80  
 SQ SEQUENCE 80 AA; 8901 MW; A4D5FCB8EE0CD3FF CRC64;

Query Match 100.0%; Score 26; DB 7; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 ||||  
 Db 7 RGDC 10

## RESULT 12

Q9XRD7 PRELIMINARY; PRT; 80 AA.  
 AC Q9XRD7;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE MHC class I alpha 2 antigen (Fragment).  
 OS Salmo salar (Atlantic salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
 OX NCBI\_TaxID=8030;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sasa-UA\*7; TISSUE=Fin;  
 RA Miller K.M., Withler R.E.;  
 RT "The salmonid class I MHC: limited diversity in a primitive teleost.";  
 RL Immunol. Rev. 0:0-0(1999).  
 DR EMBL; AF104613; AAD26109.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR001039; MHC\_I.  
 DR Pfam; PF00129; MHC\_I; 1.  
 DR PRINTS; PR01638; MHCCLASSI.  
 DR PRODOM; PD000050; MHC\_I; 1.  
 FT NON\_TER 1  
 FT NON\_TER 80  
 SQ SEQUENCE 80 AA; 8957 MW; A0D687FB2693FFC5 CRC64;

Query Match 100.0%; Score 26; DB 7; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDC 4  
 ||||  
 Db 8 RGDC 11

## RESULT 13

Q9XRD3



```

ID Q9XRD3 PRELIMINARY; PRT; 80 AA.
AC Q9XRD3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MHC class I alpha 2 antigen (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sasa-UA*1; TISSUE=Fin;
RA Miller K.M., Withler R.E.;
RT "The salmonid class I MHC: limited diversity in a primitive teleost.";
RL Immunol. Rev. 0:0-0(1999).
DR EMBL; AF104609; AAD26105.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9046 MW; 9610DF263B0FDB4 CRC64;

Query Match 100.0%; Score 26; DB 7; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 8 RGDC 11

RESULT 14
Q9XRD6 PRELIMINARY; PRT; 90 AA.
AC Q9XRD6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MHC class I alpha 2 antigen (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sasa-UA*6; TISSUE=Fin;
RA Miller K.M., Withler R.E.;
RT "The salmonid class I MHC: limited diversity in a primitive teleost.";
RL Immunol. Rev. 0:0-0(1999).
DR EMBL; AF104612; AAD26108.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 90 AA; 8903 MW; 745B17FB255F2497 CRC64;

Query Match 100.0%; Score 26; DB 7; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 8 RGDC 11

RESULT 15
Q7WXT6 PRELIMINARY; PRT; 81 AA.
AC Q7WXT6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase.
GN PHG030.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H16;
RA Schwartz E., Henne A., Cramm R., Eitinger T., Friedrich B.,
RA Gottschalk G.;
RT "Complete Nucleotide Sequence of pHG:: A Ralstonia eutropha H16
RT Megaplasmid Encoding Key Enzymes of H2-Based Lithoautotrophy and
RT Anaerobiosis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY305378; AAP85783.1; -.
KW Plasmid.
SQ SEQUENCE 81 AA; 9348 MW; 6C355D483FD282F1 CRC64;

Query Match 100.0%; Score 26; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDC 4
Db 8 RGDC 11

Search completed: April 16, 2004, 07:27:39
Job time : 34.6 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 31.6 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-2

Perfect score: 21

Sequence: 1 RGVV 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	28	4 Q9NQR3	Q9nqr3 homo sapien
2	21	100.0	28	4 Q81ZT7	Q81zt7 homo sapien
3	21	100.0	28	4 Q81ZK2	Q81zk2 homo sapien
4	21	100.0	30	5 P822C5	P822c5 bombyx mori
5	21	100.0	35	4 Q9Y634	Q9y634 homo sapien
6	21	100.0	38	9 Q857D6	Q857d6 mycobacteri
7	21	100.0	40	17 Q8ZMJ2	Q8zj2 pyrobaculum
8	21	100.0	42	17 Q8ZX76	Q8zx76 pyrobaculum
9	21	100.0	45	10 Q43718	Q43718 arabidopsis
10	21	100.0	48	16 Q7UND9	Q7und9 rhodospirell
11	21	100.0	51	4 Q9Y246	Q9y246 homo sapien
12	21	100.0	56	12 Q93056	Q93056 pseudorabie
13	21	100.0	56	17 Q9HNT0	Q9hnt0 halobacteri
14	21	100.0	58	5 Q2C348	Q2c348 caenorhabdi
15	21	100.0	61	10 Q9C8M4	Q9c8m4 arabidopsis
16	21	100.0	62	5 Q25250	Q25250 lucilia cup

17	21	100.0	62	12 Q91D93	Q91d93 hepatitis c
18	21	100.0	62	16 Q83NE9	Q83ne9 tropheryma
19	21	100.0	62	16 Q83GQ7	Q83gq7 tropheryma
20	21	100.0	62	16 Q7VER3	Q7ver3 mycobacteri
21	21	100.0	65	2 Q44842	Q44842 bartonella
22	21	100.0	65	10 Q43851	Q43851 trititicum ae
23	21	100.0	66	5 Q95VA9	Q95va9 spirometra
24	21	100.0	66	16 Q8XVQ8	Q8xvq8 ralstonia s
25	21	100.0	67	10 Q8RZ72	Q8rz72 oryza sativ
26	21	100.0	68	2 Q9RC06	Q9rc06 sphingomona
27	21	100.0	69	17 Q8TXX3	Q8txt3 methanopyru
28	21	100.0	71	9 Q85618	Q85618 mycobacteri
29	21	100.0	71	16 Q8XZQ3	Q8xzq3 ralstonia s
30	21	100.0	73	5 Q76817	Q76817 anopheles g
31	21	100.0	75	15 Q99176	Q991t6 equine infe
32	21	100.0	75	16 Q835D9	Q835d9 enterococcu
33	21	100.0	75	17 Q29645	Q29645 archaeoglob
34	21	100.0	76	16 Q836W5	Q836w5 enterococcu
35	21	100.0	77	16 Q9KLU5	Q9klm5 vibrio chol
36	21	100.0	77	15 Q9K477	Q9k477 streptomyce
37	21	100.0	78	2 Q9EU09	Q9eu09 bacteroides
38	21	100.0	78	16 Q8KAU8	Q8kau8 chlorobium
39	21	100.0	79	7 Q9XRC2	Q9xrc2 oncorhynch
40	21	100.0	79	7 Q9TNY8	Q9tny8 oncorhynch
41	21	100.0	79	17 Q28591	Q28591 archaeoglob
42	21	100.0	80	5 Q03624	Q03624 plasmodium
43	21	100.0	80	7 P79647	P79647 oncorhynch
44	21	100.0	80	16 Q82NQ1	Q82nq1 streptomyce
45	21	100.0	82	16 Q8KF71	Q8kf71 chlorobium

## ALIGNMENTS

### RESULT 1

Q9NQR3	PRELIMINARY;	PRT;	28 AA.
ID	Q9NQR3		
AC	Q9NQR3;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	BRCA1 (Breast and ovarian cancer susceptibility protein)		
DE	(Fragment).		
GN	BRCA1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yassae V.R.;		
RT	"Novel Mutation in Iranian women with Breast Cancer.";		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[6]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[7]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[8]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[9]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[10]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[11]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[12]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[13]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[14]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[15]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[16]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[17]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[18]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[19]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[20]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[21]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[22]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[23]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[24]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[25]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[26]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[27]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[28]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[29]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[30]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[31]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[32]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[33]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[34]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[35]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[36]		
RP	SEQUENCE FROM N.A.		
RA	Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;		
RT	"BRCA1 germline mutations in Indian breast cancer families.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[37]		

```

FT NON TER 28 28
SQ SEQUENCE 28 AA; 3289 MW; 59863AA07B9F445E CRC64;

Query Match 100.0%; Score 21; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4
Db 6 RGDV 9

RESULT 2
Q81ZT7 PRELIMINARY; PRT; 28 AA.
AC Q81ZT7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Breast and ovarian cancer susceptibility protein (Fragment).
GN BCAL
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;
RT "BRCA1 germline mutations in Indian breast cancer families.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507077; AA61424.1; -.
FT NON TER 1
FT NON TER 28
SQ SEQUENCE 28 AA; 3220 MW; 598633D90B9F445E CRC64;

Query Match 100.0%; Score 21; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4
Db 6 RGDV 9

RESULT 3
Q81ZK2 PRELIMINARY; PRT; 28 AA.
AC Q81ZK2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Breast and ovarian cancer susceptibility protein (Fragment).
GN BCAL
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Valarmathi M.T., Agarwal A., Deo S.S.V., Shukla N.K., Das S.N.;
RT "BRCA1 germline mutations in Indian breast cancer families.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093492; AAM18225.1; -.
FT NON TER 1
FT NON TER 28
SQ SEQUENCE 28 AA; 3228 MW; 59926BES7B9F445E CRC64;

Query Match 100.0%; Score 21; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4
Db 6 RGDV 9

```

```

Db 6 RGDV 9

RESULT 4
P82205 PRELIMINARY; PRT; 30 AA.
AC P82205;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Superoxide dismutase (EC 1.15.1.1) (Fragment).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE.
RC STRAIN=XINHANG X KEMING;
RC TISSUE=Fifth instar larvae body wall, and Fifth instar larvae fat
RC body;
RX MEDLINE=21177481; PubMed=11280994;
RA Zhong B.X.;
RT "Protein database for several tissues derived from five instar of
RT silkworm.";
RL I Chuan Hsueh Pao 28:217-224(2001).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0008382; F:iron superoxide dismutase activity; IEA.
DR GO; GO:0008383; F:manganese superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:nickel ion binding; IEA.
DR GO; GO:0016954; F:nickel superoxide dismutase activity; IEA.
DR GO; GO:0016431; F:oxidoreductase activity; IEA.
DR GO; GO:0006801; F:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sdcu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; PARTIAL.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Oxidoreductase.
FT NON_TER 30
FT NON_TER 30
SQ SEQUENCE 30 AA; 3179 MW; CB96A9EF97338940 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4
Db 9 RGDV 12

RESULT 5
Q9Y634 PRELIMINARY; PRT; 35 AA.
AC Q9Y634;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Calcium/calmodulin-stimulated cyclic nucleotide phosphodiesterase
DE (Fragment).
DE PDE1A.
GN PDE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;

```

RX MEDLINE=99334236; PubMed=10405764;  
 RA Snyder P.B., Florio V.A., Ferguson K., Loughney K.;  
 RT "Isolation, expression and analysis of splice variants of a human  
 RI Ca2+/calmodulin-stimulated phosphodiesterase (PDE1A).";  
 RJ Cell. Signal. 11:535-544(1999).  
 DR EMBL; AF110235; AAD40737.1; -.  
 FT NON TER 35 35  
 SQ SEQUENCE 35 AA; 4229 MW; 14877AD07AAEFEBE CRC64;

Query Match 100.0%; Score 21; DB 4; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 28 RGDV 31

RESULT 6  
 Q857D6 PRELIMINARY; PRT; 38 AA.  
 ID Q857D6  
 AC Q857D6  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gb86.  
 OS Mycobacteriophage Bx22.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=205870;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22592660; PubMed=12705866;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
 RA Brucker W., Kumar V., Kandasamy S., Keenan L., Bardarov S.,  
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
 RA Hatfull G.F.;  
 RT "Origins of highly mosaic mycobacteriophage genomes.";  
 RL Cell 113:171-182(2003).  
 DR EMBL; AY129332; AAN01840.1; -.  
 SQ SEQUENCE 38 AA; 4363 MW; 7C35E568C7BC9972 CRC64;

Query Match 100.0%; Score 21; DB 9; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 14 RGDV 17

RESULT 7  
 Q82WJ2 PRELIMINARY; PRT; 40 AA.  
 ID Q82WJ2  
 AC Q82WJ2  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE PaREP2a.  
 GN PAE1765A.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 CC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX MEDLINE=21664397; PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

DR EMBL; AE009838; AAL63710.1; -.  
 XW Complete proteome.  
 SQ SEQUENCE 40 AA; 4439 MW; CCF05935688D1365 CRC64;

Query Match 100.0%; Score 21; DB 17; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 12 RGDV 15

RESULT 8  
 Q8ZX76 PRELIMINARY; PRT; 42 AA.  
 ID Q8ZX76  
 AC Q8ZX76  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein PAE1420.  
 GN PAE1420.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 CC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX MEDLINE=21664397; PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 DR EMBL; AE009821; AAL63473.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 42 AA; 5055 MW; A812C27D4984BD0 CRC64;

Query Match 100.0%; Score 21; DB 17; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 33 RGDV 36

RESULT 9  
 O49718 PRELIMINARY; PRT; 45 AA.  
 ID O49718  
 AC O49718  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN T805.160 OR A14G21950.  
 CS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Wedler H., Mambutt R., Bancroft I., Mewes H.W., Mayer K.,  
 RA Schueller C.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Wedler H., Mambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021890; CAA17162.1; -.  
 DR EMBL; AL161556; CAB79150.1; -.

```

DR PIR; T05477; T05477.
KW Hypothetical protein.
SQ SEQUENCE 45 AA; 5177 MW; F48ABE0DC143A851 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 25 RGDV 28

RESULT 13
Q7UND9 PRELIMINARY; PRT; 48 AA.
AC Q7UND9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB7632.
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Firellula.
OX NCBI_TaxID=1117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=27735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Firellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
RL EMBL; BX254146; CAD75480.1; -.
DR EMBL; BX254146; CAD75480.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 48 AA; 5350 MW; 6D57E475E50CD989 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 33 RGDV 36

RESULT 11
Q9Y246 PRELIMINARY; PRT; 51 AA.
AC Q9Y246;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-2001 (TrEMBLrel. 16, Last annotation update)
DE Calcium/calmodulin-stimulated cyclic nucleotide phosphodiesterase
DE (Fragment).
GN PDE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex, and lung;
RX MEDLINE=99334236; PubMed=10405764;
RA Snyder P.B., Florio V.A., Ferguson K., Loughney K.;
RT "Isolation, expression and analysis of splice variants of a human
RT Ca2+/calmodulin-stimulated phosphodiesterase (PDE1A).";
RL Cell. Signal. 11:535-544 (1999).
RL EMBL; AF110237; AAC40739.1; -.
DR EMBL; AF110237; AAC40739.1; -.
DR EMBL; AF110236; AAC40738.1; -.

FT NON_TER 51 51
SQ SEQUENCE 51 AA; 5889 MW; 3599B22E2A9CDF45 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 44 RGDV 47

RESULT 12.
Q93056 PRELIMINARY; PRT; 56 AA.
AC Q93056;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA, the junction of the unique short and short terminal repeat
DE regions (Fragment).
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=33703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91037977; PubMed=2172457;
RA Zhang G., Leader D.P.;
RT "The structure of the pseudorabies virus genome at the end of inverted
RT repeat sequences proximal to the junction with the short unique
RT region."
RL J. Gen. Virol. 71:2433-2441 (1990).
RL EMBL; D00633; BAA00526.1; -.
FT NON_TER 1
SQ SEQUENCE 56 AA; 6435 MW; 168A82669EBC3881 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 33 RGDV 36

RESULT 13
Q9HNT0 PRELIMINARY; PRT; 56 AA.
AC Q9HNT0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Vag1960h.
GN VNG1960H.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L.,
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
RL EMBL; AE005092; AAG20140.1; -.

```

DR PIR: H84346; H84346.  
 KW Complete proteome.  
 SQ SEQUENCE 56 AA; 6670 MW; 5A3B3A5898792543 CRC64;  
 Query Match 100.0%; Score 21; DB 17; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 Db 10 RGDV 13  
 RESULT 14  
 Q20348 PRELIMINARY; PRT; 58 AA.  
 ID Q20348  
 AC Q20348  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN F42G9.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2.  
 RA R. Wilson., Taich A.;  
 RT "The sequence of C. elegans cosmid F42G9.";  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U00051; AAA91351.2; -.  
 DR PIR: T16347; T16347.  
 DR WormPep: F42G9.4; CE30789.  
 KW Hypothetical protein.  
 SQ SEQUENCE 58 AA; 6812 MW; 7DDC8EE7B0C91B17 CRC64;  
 Query Match 100.0%; Score 21; DB 5; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 Db 33 RGDV 36  
 RESULT 15  
 Q9C8M4 PRELIMINARY; PRT; 61 AA.  
 ID Q9C8M4  
 AC Q9C8M4;  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE RNA polymerase II, putative.  
 GN F22G10.4.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 Buehler E., Chao A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,  
 Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 DR EMBL: AC024260; AAC51987.1; -.  
 DR PIR: B96577; B96577.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0003899; F:DNA-directed RNA polymerase activity; IEA.  
 DR GO: GO:0006350; P:transcription; IEA.  
 DR InterPro: IPR003221; DNA\_RNAPol\_7KD.  
 DR InterPro: IPR006591; DNA\_pol\_Rbp10.  
 DR Pfam: PF03604; DNA\_RNAPol\_7kd; 1.  
 DR ProDom: PD012151; DNA\_RNAPol\_7kd; 1.  
 DR SMART: SM00659; RPOLCX; 1.  
 SQ SEQUENCE 61 AA; 7137 MW; 46A4AB41F797FD53 CRC64;  
 Query Match 100.0%; Score 21; DB 10; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDV 4  
 Db 32 RGDV 35

Search completed: April 16, 2004, 07:27:42  
 Job time : 34.6 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: April 16, 2004, 07:28:08 ; Search time 26.9 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588b-2

Perfect score: 21

Sequence: 1 RGDV 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 3%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	4	20	US-09-911-569-16
2	21	100.0	4	11	US-09-991-588b-2
3	21	100.0	4	14	US-10-215-435-7
4	21	100.0	4	14	US-10-200-879-16
5	21	100.0	4	15	US-10-297-229-41
6	21	100.0	10	12	US-09-995-388-46
7	21	100.0	11	12	US-10-033-769-9
8	21	100.0	13	14	US-10-116-519-73
9	21	100.0	13	14	US-10-116-519-89
10	21	100.0	15	12	US-10-402-212-8
11	21	100.0	15	14	US-10-115-223-8
12	21	100.0	16	14	US-10-225-567A-1361
13	21	100.0	20	10	US-09-962-756-1062
14	21	100.0	20	14	US-10-280-066-469
15	21	100.0	20	15	US-10-253-471-1062

16	21	100.0	20	16	US-10-253-493-1062
17	21	100.0	21	9	US-09-782-977-1
18	21	100.0	21	9	US-09-812-485A-13
19	21	100.0	24	13	US-10-016-157A-232
20	21	100.0	25	10	US-09-986-480-406
21	21	100.0	29	14	US-10-121-282-39
22	21	100.0	29	14	US-10-105-232-512
23	21	100.0	29	14	US-10-189-437-499
24	21	100.0	30	14	US-10-105-232-466
25	21	100.0	30	14	US-10-189-437-453
26	21	100.0	32	14	US-10-105-232-511
27	21	100.0	32	14	US-10-189-437-498
28	21	100.0	33	14	US-10-105-232-465
29	21	100.0	33	14	US-10-189-437-452
30	21	100.0	34	9	US-09-864-761-47266
31	21	100.0	34	14	US-10-105-232-520
32	21	100.0	34	14	US-10-189-437-507
33	21	100.0	35	14	US-10-105-232-510
34	21	100.0	35	14	US-10-189-437-497
35	21	100.0	36	14	US-10-105-232-464
36	21	100.0	36	14	US-10-189-437-451
37	21	100.0	37	14	US-10-105-232-519
38	21	100.0	37	14	US-10-189-437-506
39	21	100.0	38	14	US-10-105-232-516
40	21	100.0	38	14	US-10-189-437-503
41	21	100.0	39	14	US-10-105-232-508
42	21	100.0	39	14	US-10-189-437-495
43	21	100.0	40	14	US-10-153-344-25
44	21	100.0	40	14	US-10-105-232-461
45	21	100.0	40	14	US-10-105-232-463

## ALIGNMENTS

### RESULT 1

US-09-911-569-16

; Sequence 16, Application US/09911569

; Publication No. US20030069173A1

; GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA

SHIH, FOJEN

JESSE, JOEL A.

SCHEFFERLI, KEVIN P.

GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/911.569

FILING DATE: 23-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-808C



TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-911-569-16

Query Match 100.0%; Score 21; DB 13; Length 4;  
Best Local Similarity 100.0%; Pred. No. le+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
Db 1 RGDV 4

RESULT 2  
US-09-991-588B-2  
Sequence 2, Application US/09991588B  
Publication No. US20030219429A1  
GENERAL INFORMATION:  
APPLICANT: Budhy, John A.  
TITLE OF INVENTION: Composition and Method for Bone Regeneration  
FILE REFERENCE: 1008-120 US  
CURRENT APPLICATION NUMBER: US/09/991,588B  
CURRENT FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 09/122,348  
PRIOR FILING DATE: 1998-07-24  
NUMBER OF SEQ ID NOS: 23  
SEQ ID NO 2  
LENGTH: 4  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Purchased commercially or sequence is synthesized  
US-09-991-588B-2

Query Match 100.0%; Score 21; DB 11; Length 4;  
Best Local Similarity 100.0%; Pred. No. le+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
Db 1 RGDV 4

RESULT 3  
US-10-215-435-7  
Sequence 7, Application US/10215435  
Publication No. US20030104614A1  
GENERAL INFORMATION:  
APPLICANT: Urrich, Kathryn E.  
APPLICANT: Buettner, Helen  
APPLICANT: Schmalenberg, Kristine  
TITLE OF INVENTION: Micropatterning Surfaces of Polymeric Substrates  
FILE REFERENCE: 1435.009US  
CURRENT APPLICATION NUMBER: US/10/215,435  
CURRENT FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: PCT/US01/04842  
PRIOR FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/181,763  
PRIOR FILING DATE: 2000-02-11  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 4  
TYPE: PRT

ORGANISM: Unknown;  
FEATURE:  
OTHER INFORMATION: A peptide sequence found in vitronectin.  
US-10-215-435-7

Query Match 100.0%; Score 21; DB 14; Length 4;  
Best Local Similarity 100.0%; Pred. No. le+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
Db 1 RGDV 4

RESULT 4  
US-10-200-879-16  
Sequence 16, Application US/10200879  
Publication No. US20030144230A1  
GENERAL INFORMATION:  
APPLICANT: HAWLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GSEYEHU, GUILIAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/200,879  
FILING DATE: 23-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/911,569  
FILING DATE: 23-JUL-2001  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
APPLICATION NUMBER: US 08/818,200  
FILING DATE: 14-MAR-1997  
APPLICATION NUMBER: US 08/658,130  
FILING DATE: 04-JUN-1996  
APPLICATION NUMBER: US 08/477,354  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-10-200-879-16

Query Match 100.0%; Score 21; DB 14; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 1 RGDV 4

RESULT 5  
US-10-297-229-41  
; Sequence 41, Application US/10297229  
; Publication No. US20030220245A1  
; GENERAL INFORMATION:  
; APPLICANT: HUBBELL, Jeffrey A.  
; APPLICANT: ELBERT, Donald  
; APPLICANT: SCHOENMAKERS, Ronald  
; TITLE OF INVENTION: CONJUGATE ADDITION REACTIONS FOR THE  
; FILE REFERENCE: 50154/003002  
; CURRENT APPLICATION NUMBER: US/10/297,229  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/US01/18101  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: US 09/586,937  
; PRIOR FILING DATE: 2000-06-02  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Based on Homo sapiens  
US-10-297-229-41

Query Match 100.0%; Score 21; DB 15; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 1 RGDV 4

RESULT 6  
US-09-995-388-46  
; Sequence 46, Application US/09995388  
; Publication No. US20020106325A1  
; GENERAL INFORMATION:  
; APPLICANT: Carpenter, Jr., Alan P.  
; TITLE OF INVENTION: SIMULTANEOUS IMAGING OF CARDIAC PERFUSION AND A VITRONECTIN  
; FILE REFERENCE: BMS-2201  
; CURRENT APPLICATION NUMBER: US/09/995,388  
; CURRENT FILING DATE: 2001-11-27  
; PRIOR APPLICATION NUMBER: US 60/253,324  
; PRIOR FILING DATE: 2000-11-27  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 46  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-09-995-388-46

Query Match 100.0%; Score 21; DB 12; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4

Db 1 RGDV 4

RESULT 7  
US-10-033-769-9  
; Sequence 9, Application US/10033769  
; Publication No. US20020094316A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Shuang  
; TITLE OF INVENTION: POLYPYODAL CHELANTS FOR METALLOPHARMACEUTICALS  
; FILE REFERENCE: BMS-2204  
; CURRENT APPLICATION NUMBER: US/10/033,769  
; CURRENT FILING DATE: 2001-12-27  
; PRIOR APPLICATION NUMBER: US 60/260,619  
; PRIOR FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: misc\_feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: Xaa = BTD  
US-10-033-769-9

Query Match 100.0%; Score 21; DB 12; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 1 RGDV 4

RESULT 8  
US-10-116-519-73  
; Sequence 73, Application US/10116519  
; Publication No. US20030114373A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN  
; FILE REFERENCE: D0124 NP  
; CURRENT APPLICATION NUMBER: US/10/116,519  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: US 60/281,253  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/288,768  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/296,180  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/300,620  
; PRIOR FILING DATE: 2001-06-25  
; NUMBER OF SEQ ID NOS: 145  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 73  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-519-73

Query Match 100.0%; Score 21; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
DB 6 RGDV 9

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 RGDV 4  
Db 11 RGDV 14

RESULT 9  
US-10-116-519-89  
; Sequence 89, Application US/10116519  
; Publication No. US20030114373A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN  
; FILE REFERENCE: D0124 NP  
; CURRENT APPLICATION NUMBER: US/10/116,519  
; PRIOR FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: US 60/281,253  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/289,768  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/296,180  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/300,620  
; PRIOR FILING DATE: 2001-06-25  
; NUMBER OF SEQ ID NOS: 145  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 89  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-519-89

Query Match 100.0%; Score 21; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 RGDV 4  
Db 6 RGDV 9

RESULT 10  
US-10-402-212-8  
; Sequence 8, Application US/10402212  
; Publication No. US20040063790A1  
; GENERAL INFORMATION:  
; APPLICANT: Brooks, Peter C.  
; APPLICANT: Cheres, David A.  
; APPLICANT: Silletti, Steven A.  
; TITLE OF INVENTION: THE SCRIPPS RESEARCH INSTITUTE  
; FILE REFERENCE: TSRI-419.3  
; CURRENT APPLICATION NUMBER: US/10/402,212  
; PRIOR FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: 10/115,223  
; PRIOR FILING DATE: 2002-04-02  
; PRIOR APPLICATION NUMBER: 09/194,468  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: PCT/US97/09158  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: 60/018,773  
; PRIOR FILING DATE: 1996-05-31  
; PRIOR APPLICATION NUMBER: 60/015,869  
; PRIOR FILING DATE: 1996-05-31  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-10-402-212-8

Query Match 100.0%; Score 21; DB 12; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

US-10-115-223-8  
; Sequence 8, Application US/10115223  
; Publication No. US20030176334A1  
; GENERAL INFORMATION:  
; APPLICANT: Brooks, Peter  
; APPLICANT: Cheres, David A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF  
; FILE REFERENCE: MER0049S  
; CURRENT APPLICATION NUMBER: US/10/115,223  
; CURRENT FILING DATE: 2002-04-02  
; PRIOR APPLICATION NUMBER: US/09/194,468  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/018,773  
; PRIOR FILING DATE: 1996-05-31  
; PRIOR APPLICATION NUMBER: 60/015,896  
; PRIOR FILING DATE: 1996-05-31  
; PRIOR APPLICATION NUMBER: PCT/US97/09158  
; PRIOR FILING DATE: 1997-05-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-10-115-223-8

Query Match 100.0%; Score 21; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 RGDV 4  
Db 11 RGDV 14

RESULT 12  
US-10-225-567A-1361  
; Sequence 1361, Application US/10225567A  
; Publication No. US20030113798A1  
; GENERAL INFORMATION:  
; APPLICANT: Lifespan Biosciences  
; APPLICANT: Brown, Joseph P.  
; APPLICANT: Burner, Glenn C.  
; APPLICANT: Roush, Christine L.  
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 1920-4-4  
; CURRENT APPLICATION NUMBER: US/10/225,567A  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/257,144  
; PRIOR FILING DATE: 2000-12-19  
; NUMBER OF SEQ ID NOS: 2292  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1361  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-225-567A-1361

Query Match 100.0%; Score 21; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 RGDV 4
Db      11 RGDV 14

RESULT 13
US-09-962-756-1062
; Sequence 1062, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1062
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-1062

Query Match      100.0%; Score 21; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
Db      7 RGDV 10

RESULT 14
US-10-280-066-469
; Sequence 469, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 469
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

```
; OTHER INFORMATION: FGFR1b-20F-A1
US-10-280-066-469

Query Match      100.0%; Score 21; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
Db      3 RGDV 6

RESULT 15
US-10-253-471-1062
; Sequence 1062, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1062
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-1062

Query Match      100.0%; Score 21; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
Db      7 RGDV 10

Search completed: April 16, 2004, 08:03:16
Job time : 27.9 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 8.4 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-2

Perfect score: 21

Sequence: 1 RGDV 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: Pirl:\*

2: Pirl:\*

3: Pirl:\*

4: Pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	15	2 PQ0017	terminal protein -
2	21	100.0	44	2 S50088	superoxide dismuta
3	21	100.0	45	2 T05477	hypothetical prote
4	21	100.0	51	2 T02517	hypothetical prote
5	21	100.0	56	2 H84346	hypothetical prote
6	21	100.0	61	2 B96577	probable RNA polym
7	21	100.0	65	2 T06349	glucan endo-1,3-be
8	21	100.0	75	2 B69326	hypothetical prote
9	21	100.0	77	2 C82409	hypothetical prote
10	21	100.0	79	2 A69460	hypothetical prote
11	21	100.0	82	2 D63396	antigen 7H8/7 - ma
12	21	100.0	83	2 A13439	exodeoxyribonuclea
13	21	100.0	86	2 A26568	exodeoxyribonuclea
14	21	100.0	86	2 C97450	exodeoxyribonuclea
15	21	100.0	91	2 AB1573	carboxysome struct
16	21	100.0	91	2 AG1219	carboxysome struct
17	21	100.0	91	2 AH0760	probable propanedi
18	21	100.0	94	2 G84269	30S ribosomal prot
19	21	100.0	94	2 AD1222	probable carboxyso
20	21	100.0	94	2 AG1575	probable carboxyso
21	21	100.0	94	2 AI0759	probable propanedi
22	21	100.0	95	2 AB1572	Salmonella typhimu
23	21	100.0	95	2 AG1218	Salmonella typhimu
24	21	100.0	96	2 T03014	hypothetical prote
25	21	100.0	96	2 T28901	hypothetical prote
26	21	100.0	100	1 F70309	protein export mem
27	21	100.0	100	2 B71567	hypothetical prote
28	21	100.0	100	2 G81718	conserved hypothet
29	21	100.0	101	2 H83114	30S ribosomal prot

## ALIGNMENTS

### RESULT 1

PQ0017

terminal protein - phage M2 (fragment)

C;Species: phage M2  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999

C;Accession: PQ0017

R;Matsumoto, K.; Takao, H.; Kim, C.I.; Hirokawa, H.

Gene 84, 247-255, 1989

A;Title: Primary structure of bacteriophage M2 DNA polymerase: conserved segments within

A;Reference number: JQ0161; MUID:90128268; PMID:2515115

A;Accession: PQ0017

A;Molecule type: DNA

A;Residues: 1-15 <MAT>

A;Cross-references: GB:M33144; NID:g215507; PIDN:AAA32367.1; PID:g215508

C;Genetics:

C;Gene: E

C;Superfamily: phage PZA terminal protein

Query Match 100.0%; Score 21; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4

Db 5 RGDV 8

### RESULT 2

S50088

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) 1 - Caenorhabditis briggsae (fragment)

C;Species: Caenorhabditis briggsae

C;Date: 14-Jul-1995 #sequence\_revision 01-Nov-1996 #text\_change 17-Mar-1999

C;Accession: S50088

R;Kuwabara, P.E.; Shah, S.

Nucleic Acids Res. 22, 4414-4418, 1994

A;Title: Cloning by synteny: identifying C.briggsae homologues of C.elegans genes.

A;Reference number: S50085; MUID:95061402; PMID:7971272

A;Accession: S50088

A;Molecule type: mRNA

A;Residues: 1-44 <KUM>

C;Function:

A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C;Superfamily: superoxide dismutase (Cu-Zn)

C;Keywords: metalloprotein; oxidoreductase

Query Match 100.0%; Score 21; DB 2; Length 44;

Best Local Similarity 100.0%; Pred. No. 1.8e+32;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4

Db 10 RGDV 13

## RESULT 3

T05477  
 Hypothetical protein T805.160 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999  
 C:Accession: T05477  
 R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuelle  
 submitted to the Protein Sequence Database, February 1998  
 A:Reference number: Z15417  
 A:Accession: T05477  
 A:Molecule type: DNA  
 A:Residues: 1-45 <BEV>  
 A:Cross-references: EMBL:AL021890  
 A:Experimental source: cultivar Columbia; BAC clone T805  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 11/1  
 A:Note: T805.160

Query Match 100.0%; Score 21; DB 2; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 25 RGDV 28

## RESULT 4

T02517  
 Hypothetical protein F13M22.1 - Arabidopsis thaliana (fragment)  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 14-May-1999  
 C:Accession: T02517  
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,  
 submitted to the EMBL Data Library, June 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.  
 A:Reference number: Z14677  
 A:Accession: T02517  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-51 <RCU>  
 A:Cross-references: EMBL:AC034684; NID:g3236234; PID:g3236260  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Map position: 2  
 A:Note: F13M22.1

Query Match 100.0%; Score 21; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 2 RGDV 5

## RESULT 5

H04346  
 Hypothetical protein Vcgl1960k [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: H04346  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A04160; MUID:20504483; PMID:11016950  
 A:Accession: H04346

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-56 <STO>  
 A:Cross-references: GB:AE004437; NID:gi0581397; PIDN:AG20140.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG1960H

Query Match 100.0%; Score 21; DB 2; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 10 RGDV 13

## RESULT 6

B96577  
 Probable RNA polymerase II, 28841-23486 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Aug-2002  
 C:Accession: B96577  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 ansen, K.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: B96577  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-61 <STO>  
 A:Cross-references: GB:AE005173; NID:gi0645362; PIDN:AG21482.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F22G10.4  
 A:Map position: 1  
 C:Superfamily: DNA-directed RNA polymerase chain ABC10 alpha

Query Match 100.0%; Score 21; DB 2; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4  
 Db 32 RGDV 35

## RESULT 7

T06349  
 Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) - wheat (fragment)  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 20-Jun-2000  
 C:Accession: T06349  
 R:Neubaus, J.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z15615  
 A:Accession: T06349  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-65 <NEU>  
 A:Cross-references: EMBL:X95647; NID:gi235981; PIDN:CAA64949.1  
 A:Experimental source: cv. Prelude, leaf  
 C:Superfamily: beta-1,3-glucanase  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 100.0%; Score 21; DB 2; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 RGDV 4
      19 RGDV 22
Db

RESULT 8
B69326
hypothetical protein AF0610 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: B69326
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
  A.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
  Nature 390, 364-370, 1997
  A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
  Smith, H.O.; Woese, C.R.; Venter, J.C.
  A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
  A:Reference number: A69250; MUID:98049343; PMID:9389475
  A:Accession: B69326
  A:Status: preliminary; nucleic acid sequence not shown; translation not shown
  A:Molecule type: DNA
  A:Residues: 1-75 <KLE>
  A:Cross-references: GB:AE001062; GB:AE000782; NID:g2689385; PIDN:AA90631.1; PID:g265001

      Query Match      100.0%; Score 21; DB 2; Length 75;
      Best Local Similarity 100.0%; Pred. No. 3.1e+02;
      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
      49 RGDV 52
Db

RESULT 9
C82409
hypothetical protein VCA0861 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82409
R:Reid, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
  Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.
  Nature 406, 477-483, 2003
  A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
  A:Reference number: A82335; MUID:20406833; PMID:10952301
  A:Accession: C82409
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-77 <HPI>
  A:Cross-references: GB:AE004413; GB:AE003853; NID:g9458280; PIDN:AAF96759.1; GSPDB:GN001
  A:Experimental source: serogroup O1; strain N16961; biotype El Tor
  C:Genetics:
  A:Gene: VCA0861
  A:Map position: 2

      Query Match      100.0%; Score 21; DB 2; Length 77;
      Best Local Similarity 100.0%; Pred. No. 3.2e+02;
      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
      37 RGDV 40
Db

RESULT 10
A69460
hypothetical protein AF1682 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: A69460
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
  A.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
  Nature 390, 364-370, 1997
  A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
  Smith, H.O.; Woese, C.R.; Venter, J.C.
  A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
  A:Reference number: A69250; MUID:98049343; PMID:9389475
  A:Accession: A69460
  A:Status: preliminary; nucleic acid sequence not shown; translation not shown
  A:Molecule type: DNA
  A:Residues: 1-79 <KLE>
  A:Cross-references: GB:AE000987; GB:AE000782; NID:g2689310; PIDN:AA89575.1; PID:g26488E

      Query Match      100.0%; Score 21; DB 2; Length 79;
      Best Local Similarity 100.0%; Pred. No. 3.3e+02;
      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
      46 RGDV 49
Db

RESULT 11
D60396
antigen 7H8/7 - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C:Accession: D60396
R:Limpalaboon, T.; Taylor, D.W.; Jones, G.; Geyse, H.M.; Saul, A.
  Southeast Asian J. Trop. Med. Public Health 21, 388-396, 1990
  A:Title: Characterization of a Plasmodium falciparum epitope recognized by a monoclonal
  A:Reference number: A60396; MUID:91164876; PMID:1706114
  A:Accession: D60396
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-82 <LIM>
  A:Cross-references: GB:M63819
  C:Superfamily: pol. polypeptide

      Query Match      100.0%; Score 21; DB 2; Length 82;
      Best Local Similarity 100.0%; Pred. No. 3.4e+02;
      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
      53 RGDV 56
Db

RESULT 12
AI3439
exodeoxyribonuclease VII (EC 3.1.11.6) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002
C:Accession: AI3439
R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
  A.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
  Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
  A:Title: the genome sequence of the facultative intracellular pathogen Brucella melitensis
  A:Reference number: A03252; PMID:11756688
  A:Accession: AI3439
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-83 <KUR>
  A:Cross-references: GB:AE008917; PIDN:AA152684.1; PID:gl7983510; GSPDB:GN00190
  A:Experimental source: strain 16M
  C:Genetics:
  A:Gene: BME11503
  A:Map position: 1

      Query Match      100.0%; Score 21; DB 2; Length 83;
      Best Local Similarity 100.0%; Pred. No. 3.4e+02;
      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 RGDV 4  
 ||||  
 Db 33 RGDV 33

## RESULT 13

AB2668  
 exodeoxyribonuclease small subunit [imported] - Agrobacterium tumefaciens (strain C58, I  
 C/Species: Agrobacterium tumefaciens  
 C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C/Accession: AE2668  
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, X.; Gordon-Kamm,  
 s-er, E.W.  
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A/Reference number: AB2577; PMID:21608550; PMID:11743193  
 A/Accession: AB2668  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-86 <RUR>  
 A/Cross-references: GB:AE008688; PIDN:AAL41763.1; PID:gi17739115; GSPDB:GN00186  
 A/Experimental source: strain C58 (Dupont)  
 C/Genetics:  
 A/Gene: xseB  
 A/Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 33 RGDV 36

## RESULT 14

C97450  
 exodeoxyribonuclease small chain (AF203981) [imported] - Agrobacterium tumefaciens (stra  
 C/Species: Agrobacterium tumefaciens  
 C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C/Accession: C97450  
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A/Reference number: A37359; PMID:21608551; PMID:11743194  
 A/Accession: C97450  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-86 <RUR>  
 A/Cross-references: GB:AE007869; PIDN:AAK86556.1; PID:gi15155718; GSPDB:GN00169  
 C/Genetics:  
 A/Gene: AGR\_C1356  
 A/Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 33 RGDV 36

## RESULT 15

AB1573  
 carboxysome structural protein homolog lin123 [imported] - Listeria innocua (strain Cli  
 C/Species: Listeria innocua  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 11-Aug-2003  
 C/Accession: AB1573

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloescke  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
 A/Title: Comparative genomics of Listeria species.  
 A/Reference number: AB:077; PMID:21537279; PMID:11679669  
 A/Accession: AB1573  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-91 <GLA>  
 A/Cross-references: GB:AL592022; PIDN:CAC96354.1; PID:gi16413582; GSPDB:GN00178  
 A/Experimental source: strain Clipl1262  
 C/Genetics:  
 A/Gene: lin123  
 C/Superfamily: carboxysome/polyhedral organelle shell protein, CcmK/EutK/PduA type

Query Match 100.0%; Score 21; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 47 RGDV 50

Search completed: April 16, 2004, 07:29:17  
 Job time : 9.4 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time: 7 Seconds  
(without alignments)  
29.754 Million cell updates/sec

Title: US-09-991-588B-2

Perfect score: 21  
Sequence: 1 RGDV 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	15	1 TERM_BPM2	P19897 bacteriophage
2	21	100.0	83	1 EX7S_BRUME	O8yfl7 bruceella me
3	21	100.0	83	1 EX7S_RHILO	Q985y6 rhizobium l
4	21	100.0	83	1 EX7S_RHIME	Q92r19 rhizobium m
5	21	100.0	86	1 EX7S_AGRTS	Q8uhds agrobacteri
6	21	100.0	94	1 PDUA_SALTY	P37448 salmonella
7	21	100.0	96	1 EUTM_SALTY	P41791 salmonella
8	21	100.0	97	1 EUTM_ECOLI	P77606 escherichia
9	21	100.0	100	1 SECG_AQUAE	O66505 aquifex ae
10	21	100.0	100	1 URE3_RALSO	Q8xx88 ralstonia s
11	21	100.0	100	1 Y031_CHLTR	O84034 chlamydia t
12	21	100.0	100	1 Y300_CHLMC	Q9pl08 chlamydia m
13	21	100.0	101	1 RS14_PSEAE	Q9hwe8 pseudomonas
14	21	100.0	101	1 RS14_XVLPF	O9pe63 xylella fas
15	21	100.0	102	1 CCK2_SYNY3	P72761 synecocyst
16	21	100.0	102	1 CCKM_SYNY7	Q03511 synecococc
17	21	100.0	102	1 VATF_DESSY	O06503 desulfuroco
18	21	100.0	106	1 HUPF_RHOCA	Q03005 rhodobacter
19	21	100.0	106	1 RL36_DAUCA	P52866 daucus caro
20	21	100.0	109	1 SSI_STRAN	P01007 streptomyc
21	21	100.0	110	1 CCK1_SYNY3	P72760 synecocyst
22	21	100.0	111	1 CCK4_SYNY3	P73407 synecocyst
23	21	100.0	112	1 CCKM_SYNY2	Q10709 synecococc
24	21	100.0	113	1 RL19_COREF	Q8fp56 corynebacte
25	21	100.0	113	1 RL19_CORGL	Q8nn20 corynebacte
26	21	100.0	113	1 RL19_MYCLE	C33020 mycobacteri
27	21	100.0	113	1 RL19_MYCTU	Q10792 mycobacteri
28	21	100.0	115	1 RL19_COMBU	Q83e85 coxiella bu
29	21	100.0	116	1 GLB_PARCA	P15160 paramecium
30	21	100.0	116	1 RL19_PSEAK	Q9hxd2 pseudomonas
31	21	100.0	116	1 RL19_PSEPK	Q88mv3 pseudomonas
32	21	100.0	116	1 RL19_PSESM	Q886u9 pseudomonas
33	21	100.0	116	1 RL19_STRAW	Q82jw4 streptomyc

34	21	100.0	116	1 RL19_STRCO	O69883 streptomyc
35	21	100.0	119	1 YIQ5_YEAST	P40446 saccharomyc
36	21	100.0	121	1 RL19_NEIMA	Q9jvl1 neisseria m
37	21	100.0	121	1 RL19_NEIMB	Q9k0k5 neisseria m
38	21	100.0	122	1 RL14_MYCLE	O32993 mycobacteri
39	21	100.0	122	1 RL14_MYCTU	P95062 mycobacteri
40	21	100.0	122	1 RL19_CHRPU	Q7nr77 chromobacte
41	21	100.0	124	1 RL22_TREPA	O83224 treponema p
42	21	100.0	126	1 RL19_BORBP	Q7wgt8 bordetella
43	21	100.0	126	1 RL19_BORPA	Q7vze0 bordetella
44	21	100.0	126	1 RL19_BORPE	Q8y0v7 ralstonia s
45	21	100.0	128	1 RL19_RALSO	

#### ALIGNMENTS

RESULT 1  
TERM\_BPM2  
ID TERM BPM2 STANDARD; PRT; 15 AA.  
AC P19897;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA terminal protein (Protein GP3) (Fragment).  
GN 3 OR E.  
OS Bacteriophage M2.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC phi-29-like viruses.  
OX NCBI\_TaxID=10751;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90128268; PubMed=2515115;  
RA Matsumoto K., Takano H., Kim C.I., Hirokawa H.;  
RT "Primary structure of bacteriophage M2 DNA polymerase: conserved  
RT segments within protein-priming DNA polymerases and DNA polymerase I  
RT of Escherichia coli";  
RL Gene 84:247-255(1989).  
CC -!- FUNCTION: DNA terminal protein is linked to the 5' ends of both  
CC strands of the genome through a phosphodiester bond between the  
CC beta-hydroxyl group of a serine residue and the 5'-phosphate of  
CC the terminal deoxyadenylate. This protein is essential for DNA  
CC replication and is involved in the priming of DNA elongation.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; M33144; AAA32367.1; --  
CC PIR; PQ0017; PQ0017.  
CC InterPro; IPR008770; Phi-29\_GP3.  
CC Pfam; PF05435; Phi-29\_GP3; I.  
CC Early protein; DNA replication; DNA priming;  
CC Covalent protein-DNA linkage.  
CC NON TER 1 1  
CC SITE 5 7 CELL ATTACHMENT SITE (POTENTIAL).  
CC SEQUENCE 15 AA; 1797 MW; D3CBAPF8759DERA06 CRC64;  
DR PIR; PQ0017; PQ0017.  
DR InterPro; IPR008770; Phi-29\_GP3.  
DR Pfam; PF05435; Phi-29\_GP3; I.  
DR Early protein; DNA replication; DNA priming;  
DR Covalent protein-DNA linkage.  
DR NON TER 1 1  
DR SITE 5 7 CELL ATTACHMENT SITE (POTENTIAL).  
DR SEQUENCE 15 AA; 1797 MW; D3CBAPF8759DERA06 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDV 4

Db 5 RGDV 8

RESULT 2  
EX7S\_BRUME

```

ID EX7S BRUME STANDARD; PRT; 83 AA.
AC Q8VFL7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (Exonuclease VII small subunit).
GN XSEB OR BME11503 OR BR0431.
OS Brucella melitensis, and
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459, 29461;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=2320109; PubMed=1175668;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Cablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haelkorn R., Kyrpides N., Overbeek R.;
RA "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B.suis; STRAIN=1330 / Biovar 1;
RX MEDLINE=2247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Mizek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.R., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RA "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009586; AA052684.1; -.
CC F01; AL3439; A13439.
CC TIGR; BR0431; -.
CC HAMAP; MF 00337; -.
CC InterPro; IPR003761; Exonuc VII S.
CC Pfam; PF02609; Exonuc VII S; 1.
CC ProDom; PD028235; Exonuc VII S; 1.
CC TIGRFAMs; TIGR01280; xseB; 1.
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 83 AA; 9322 MW; 8B51C2BA8F9FF1E CRC64;
Query Match 100.0%; Score 21; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RGDV 4

```

```

Db 30 RGDV 33
|||||
RESULT 3
EX7S RHILQ STANDARD; PRT; 83 AA.
ID AC Q985Y6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (Exonuclease VII small subunit).
GN XSEB OR XSR7470.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099; PubMed=11214968;
RX MEDLINE=21082930;
RA Kureko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003011; BAB53567.1; ALT_INIT.
CC HAMAP; MF 00337; -.
CC InterPro; IPR003761; Exonuc VII S.
CC Pfam; PF02609; Exonuc VII S; 1.
CC ProDom; PD028235; Exonuc VII S; 1.
CC TIGRFAMs; TIGR01280; xseB; 1.
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 83 AA; 9239 MW; 03147DF752BDB8CF CRC64;
Query Match 100.0%; Score 21; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RGDV 4
|||||
Db 30 RGDV 33
|||||
RESULT 4
EX7S RHIME STANDARD; PRT; 83 AA.
ID AC Q92R19;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (Exonuclease VII small subunit).
GN XSEB OR R03882 OR SMC00970.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Eubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Routhy M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591785; CAC45454.1; -.
DR HAVAP; MF_00337; -.
DR InterPro; IPR003761; Exonuc VII_S.
DR Pfam; PF02609; Exonuc VII_S; 1.
DR ProDom; PD028235; Exonuc VII_S; 1.
DR TIGRfam; TIGR01280; xseB; 1.
KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 83 AA; 9259 MW; EC7F8E654F0871EC CRC64;

Query Match 100.0%; Score 21; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 30 RGDV 33

RESULT 5
EXS_AGR5 EXS_AGR5 STANDARD; PRT; 86 AA.
AC Q8UHD5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (Exonuclease VII small subunit).
GN XSEB OR ATU0747 OR AGR_C1356.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,

```

```

RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sk., C.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayvin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009042; AAL41763.1; -.
DR EMBL; AE008008; AAK86556.1; -.
DR PIR; AE2668; AE2668.
DR PIR; C97450; C97450.
DR HAMAP; MF_00337; -.
DR InterPro; IPR003761; Exonuc VII_S.
DR Pfam; PF02609; Exonuc VII_S; 1.
DR ProDom; PD028235; Exonuc VII_S; 1.
DR TIGRfam; TIGR01280; xseB; 1.
KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 86 AA; 9476 MW; B4A5CA90D276D5D5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4
DB 33 RGDV 36

RESULT 6
PDUA_SALTY PDUA_SALTY STANDARD; PRT; 94 AA.
AC E37448;
DT 01-OCT-1994 (Rel. 30, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Propanediol utilization protein pduA.
GN PDUA OR STM2038 OR STV2243 OR T0836.
OS Salmonella typhimurium, and

```

OS Salmonella typhi.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=602, 601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2;  
 RX MEDLINE=94350831; PubMed=8071226;  
 RA Chen P., Anderson D.I., Roth J.R.;  
 RT "The control region of the pdu/cob regulon in Salmonella  
 typhimurium";  
 RL J. Bacteriol. 176:5474-5482(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2;  
 RX MEDLINE=98312959; PubMed=9352910;  
 RA Bobik T.A., Xu Y., Jeter R.M., Otto K.E., Roth J.R.;  
 RT "Propanediol utilization genes (pdu) of Salmonella typhimurium: three  
 genes for the propanediol dehydratase";  
 RL J. Bacteriol. 179:6633-6639(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND REVISIONS TO C-TERMINUS.  
 RC SPECIES=S.typhimurium; STRAIN=LT2;  
 RX MEDLINE=9429843; PubMed=10498708;  
 RA Bobik T.A., Havemann G.D., Busch R.J., Williams D.S., Aldrich H.C.;  
 RT "The propanediol utilization (pdu) operon of Salmonella enterica  
 serovar typhimurium L2 includes genes necessary for formation of  
 polyhedral organelles involved in coenzyme B[12]-dependent 1,  
 2-propanediol degradation";  
 RL J. Bacteriol. 181:5967-5975(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of salmonella enterica serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhi; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Holden M.T.G., Sebaihia M.,  
 Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 Whitehead S., Barrett B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 enterica serovar Typhi CT18";  
 RL Nature 413:848-852(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhi; STRAIN=TY2 / ATCC 700931;  
 RX MEDLINE=22531367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;  
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
 and CT18";  
 RL J. Bacteriol. 185:2330-2337(2003).  
 CC -!- INDUCTION: By propanediol.  
 CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein  
 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF026270; AAB84107.2; -  
 CC EMBL; AE008790; AAL20942.1; -  
 CC EMBL; AL627273; CAD02399.1; -  
 CC EMBL; AE016937; AAO68524.1; -  
 CC DR StyGene; SG10278; pduA.  
 CC DR InterPro; IPR000249; Bact\_microcomp.  
 CC DR Pfam; PF00936; Bact\_microcomp; 1.  
 CC DR ProDom; PD003442; Bact\_microcomp; 1.  
 CC DR PROSITE; PS01139; BACT\_MICROCOMP; 1.  
 CC KW Complete proteome.  
 CC SQ SEQUENCE 94 AA; 9592 MW; 5EDC7C793F19DE1E CRC64;  
 CC  
 CC Query Match 100.0%; Score 21; DB 1; Length 94;  
 CC Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 RGVV 4  
 CC DB 48 RGVV 51  
 CC  
 CC RESULT 7  
 CC EUTM\_SALTY  
 CC ID EUTM\_SALTY STANDARD; PRT; 96 AA.  
 CC AC P41791;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE Ethanolamine utilization protein eutM precursor.  
 CC GN EUTM OR CCHA OR STW2465.  
 CC OS Salmonella typhimurium.  
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC OC Enterobacteriaceae; Salmonella.  
 CC OX NCBI\_TaxID=602;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=ATCC 14028S;  
 CC RX MEDLINE=95173114; PubMed=7868611;  
 CC RA Stojiljkovic I., Baemler A.J., Heffron F.;  
 CC RT "Ethanolamine utilization in Salmonella typhimurium: nucleotide  
 CC sequence, protein expression, and mutational analysis of the ccha  
 CC RT cchS eutE eutD eutG gene cluster";  
 CC RL J. Bacteriol. 177:1357-1366(1995).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=LT2;  
 CC RX MEDLINE=93395039; PubMed=10464203;  
 CC RA Kofoid E.C., Rappleye C.A., Stojiljkovic I., Roth J.R.;  
 CC RT "The 17-gene ethanolamine (eut) operon of Salmonella typhimurium  
 CC encodes five homologues of carboxysome shell proteins";  
 CC RL J. Bacteriol. 181:5317-5329(1999).  
 CC RN [3]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 CC RX MEDLINE=21534948; PubMed=11677609;  
 CC RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 CC RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 CC LT2.";  
 CC RL Nature 413:852-856(2001).  
 CC CC -!- FUNCTION: May be involved in the formation of a specific  
 CC microcompartment in the cell in which the metabolism of  
 CC potentially toxic by-products takes place.  
 CC -!- PATHWAY: Ethanolamine utilization.  
 CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein

```

CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC "extensive mosaic" structure revealed by the complete genome sequence
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18563; AAA83207.1; -
CC EMBL; AF093749; AAC78116.1; -
CC EMBL; AE008811; AAL21359.1; -
CC StyGene; SG10551; eutM.
CC InterPro; IPR000249; Bact_microcomp.
CC Pfam; PF00936; Bact_microcomp; 1.
CC ProDom; PD003442; Bact_microcomp; 1.
CC PROSITE; PS01139; BACT_MICROCOMP; 1.
CC Signal; Complete proteome.
CC SIGNAL 1 25 POTENTIAL.
CC CHAIN 26 96 ETHANOLAMINE UTILIZATION PROTEIN EUTM.
CC SEQUENCE 96 AA; 9842 MW; 01710D8ABB8CF6BC CRC64;
CC -----
CC Query Match 100.0%; Score 21; DB 1; Length 96;
CC Best Local Similarity 100.0%; Pred. No. 1.7e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 1 RGDV 4
CC DB 46 RGDV 49
CC -----
CC RESULT 8
CC EUTM_ECOLI STANDARD; PRT; 97 AA.
CC AC P77606;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Ethanolamine utilization protein eutM precursor.
CC GN EUTM OR CCHA OR B2457 OR C2982.
CC OS Escherichia coli, and
CC OS Escherichia coli O6.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=562, 217992;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=K12 / MG1655;
CC RX MEDLINE=97426617; PubMed=9278503;
CC RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC Mau B., Shao Y.;
CC RT "The complete genome sequence of Escherichia coli K-12";
CC RL Science 277:1453-1474 (1997).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=K12;
CC RX MEDLINE=97349980; PubMed=9205837;
CC RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
CC Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsubashi N.,
CC Mizobuchi K., Mori H., Nakade S., Nakamura Y., Shimamoto H.,
CC Oshima H., Oyama S., Saigo N., Sampei G., Satoh Y., Sivasubram S.,
CC Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
CC Yamagata S., Horiuchi T.;
CC RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
CC K-12 genome corresponding to 50,0-68.8 min on the linkage map and
CC analysis of its sequence features.";
CC RL DNA Res. 4:91-113 (1997).
CC [3]
CC SEQUENCE FROM N.A.
CC STRAIN=O6:H1 / CFTC73 / ATCC 700928;

```

```

RX MEDLINE=22389234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:17023-17024 (2002).
CC -!- FUNCTION: May be involved in the formation of a specific
CC microcompartment in the cell in which the metabolism of
CC potentially toxic by-products takes place.
CC -!- PATHWAY: Ethanolamine utilization.
CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC "extensive mosaic" structure revealed by the complete genome sequence
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000332; AAC75510.1; ALT INIT.
CC EMBL; D90874; BAA16335.1; ALT INIT.
CC EMBL; AE016764; AAN81432.1; ALT_INIT.
CC EcoGene; EG14187; eutM.
CC InterPro; IPR000249; Bact_microcomp.
CC Pfam; PF00936; Bact_microcomp; 1.
CC ProDom; PD003442; Bact_microcomp; 1.
CC PROSITE; PS01139; BACT_MICROCOMP; 1.
CC Signal; Complete proteome.
CC SIGNAL 1 25 POTENTIAL.
CC CHAIN 26 97 ETHANOLAMINE UTILIZATION PROTEIN EUTM.
CC SEQUENCE 97 AA; 9865 MW; 0A170FD8ABB8CF8 CRC64;
CC -----
CC Query Match 100.0%; Score 21; DB 1; Length 97;
CC Best Local Similarity 100.0%; Pred. No. 1.7e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 1 RGDV 4
CC DB 46 RGDV 49
CC -----
CC RESULT 9
CC SECG_AQUAE STANDARD; PRT; 100 AA.
CC AC O66505;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Protein-export membrane protein secg.
CC GN SECG OR AQ D98.
CC OS Aquifex aeolicus.
CC OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
CC OX NCBI_TaxID=63363;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=VF5;
CC RX MEDLINE=98196666; PubMed=9537320;
CC RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeek R., Snead M.A., Keiler M., Anjaj M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
CC RT "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus.";
CC RL Nature 392:353-358 (1998).
CC -!- FUNCTION: Involved in protein export. Participates in an early
CC event of protein translocation (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the secg family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE000673; AAC06462.1; --  
 DR PIR; F70309; F70309.  
 DR InterPro; IPR004692; SecG.  
 DR Pfam; PF03840; SecG; 1.  
 DR PRINTS; PR01651; SECSEXPORT.  
 DR TIGRFAMs; TIGR00810; secG; 1.  
 KW Protein transport; Translocation; Transmembrane; Complete proteome.  
 FT TRANSMEM 2 73 POTENTIAL.  
 FT TRANSXEM 53 22  
 SQ SEQUENCE 100 AA; 10464 MW; C7183A51B27FA21A CRC64;

Query Match 100.0%; Score 21; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 26 RGDV 29

## RESULT 10

URE3\_RALSO URE3\_RALSO STANDARD; PRT; 100 AA.  
 AC Q8XYS8;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Urease gamma subunit (EC 3.5.1.5) (Urea amidohydrolase gamma subunit).  
 GN UREA OR RSC2035 OR RS03029.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 CC Burkholderiaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,  
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Siguiet P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Souchet C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
 RL Nature 415:497-502(2002).  
 CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).  
 CC -!- SUBUNIT: (Alpha, beta, gamma)(3) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILASITY: Belongs to the urease gamma subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL646067; CAD15737.1; --  
 DR HAVAP; MF CC739; -- 1.  
 DR InterPro; IPR002026; Urease\_gamma.  
 DR Pfam; PF00547; urease\_gamma; 1.  
 DR ProDom; PD002319; Urease\_gamma; 1.  
 DR TIGRFAMs; TIGR00193; urease\_gamma; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 100 AA; 10932 MW; 7E53165E302E3FD1 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 66 RGDV 69

## RESULT 11

Y031\_CHLTR Y031\_CHLTR STANDARD; PRT; 100 AA.  
 AC O84034;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein TC031.  
 GN CT031.  
 OS Chlamydia trachomatis.  
 CC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.I., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis";  
 RL Science 282:754-759(1998).  
 CC -!- SIMILASITY: BELONGS TO THE CHLAMYDIAL CPN0121/CT031/TC0300  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE001277; AAC67621.1; --  
 DR PIR; B71567; B71567.  
 KW Hypothetical protein. Complete proteome.  
 SQ SEQUENCE 100 AA; 11683 MW; 2A12B8A53EAA4F3C CRC64;

Query Match 100.0%; Score 21; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDV 4  
 ||||  
 Db 36 RGDV 39

## RESULT 12

Y300\_CHLMU Y300\_CHLMU STANDARD; PRT; 100 AA.  
 AC Q9PL08;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein TC0300.  
 GN TC0300.  
 OS Chlamydia muridarum.  
 CC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=83560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MoPn / Nigg;  
 RX MEDLINE=20150255; PubMed=10684935;

```
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Eickley B.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."; 28:1397-1406(2000).
RL Nucleic Acids Res.
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPNC121/CT031/TC0300
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE002297; AAF39165.1; -.
DR PIR; G8178; G81718.
DR TIGR; TC0300; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 100 AA; 11545 MW; 0EDC868DCDC09185 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGVV 4
DB 36 RGVV 39

RESULT 13
RS14_PSEAE STANDARD; PRT; 101 AA.
AC Q9HWE8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 30S ribosomal protein S14.
GN RPSN OR PA4250.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OX NCBI_TaxID=287;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater V.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (By similarity).
CC -!- SIMILARITY: Belongs to the S14P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EXBL; AE004841; AAC07638.1; -.
DR PIR; H83114; H83114.
DR InterPro; IPR001209; Ribosomal S14.
DR Pfam; PF00253; Ribosomal S14; 1.
DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 101 AA; 11565 MW; B9D694FDC554902 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGVV 4
DB 90 RGVV 93

RESULT 14
RS14_XYLFA STANDARD; PRT; 101 AA.
AC Q9PEG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 30S ribosomal protein S14.
GN RPSN OR XF1165.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OX NCBI_TaxID=2371;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfy E.,
RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferri J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega P.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Paixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zacc M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (By similarity).
CC -!- SIMILARITY: Belongs to the S14P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

Search completed: April 16, 2004, 07:22:10  
Job time : 8 secs

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003950; AAF83975.1; -.
DR PIR; E82718; E82718.
DR InterPro; IPR001209; Ribosomal_S14.
DR Pfam; PF00253; Ribosomal_S14; 1
DR PROSITE; PS00527; RIBOSOMAL_S14; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 101 AA; 11553 MW; 38345DF1CAF27A7 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
Db      90 RGDV 93

RESULT 15
CCK2 SYNY3      STANDARD;      PRT;      102 AA.
AC P72761;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbon dioxide concentrating mechanism protein ccmK homolog 2.
GN CCMK2 OR SLI1028.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirotsawa M., Sugibara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki M., Naruo K.,
RA Okumura S., Shimpou S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996)
CC -!- FUNCTION: May be involved in the formation of the carboxysome, a
CC polyhedral inclusion where Rubisco is sequestered (By similarity).
CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90900; BAA16776.1; -.
DR PIR; S74624; S74624.
DR InterPro; IPR00249; Bact_microcomp.
DR Pfam; PF00936; Bact_microcomp; 1.
DR ProDom; PD003442; Bact_microcomp; 1.
DR PROSITE; PS01139; BACT_MICROCOMP; 1.
KW Multigene family; Complete proteome.
FT INIT MET 0 BY SIMILARITY.
SQ SEQUENCE 102 AA; 11003 MW; A936DCEFA1673437 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDV 4
Db      46 RGDV 49
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Computer Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 10.5 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-4

Perfect score: 27

Sequence: 1 GRGDS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96:91526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	116	2 C83492	Hypothetical prote
2	27	100.0	134	2 F75582	Hypothetical prote
3	27	100.0	149	2 AH0772	Hypothetical prote
4	27	100.0	215	2 F84889	Hypothetical prote
5	27	100.0	264	1 A40019	osteopontin precu
6	27	100.0	273	2 A28512	fibronectin - chic
7	27	100.0	278	1 JS0638	osteopontin precu
8	27	100.0	287	2 S43852	neuropeptide Pol-R
9	27	100.0	294	1 A37818	osteopontin - rat
10	27	100.0	301	2 JC5811	osteopontin precu
11	27	100.0	303	1 GEPGO	osteopontin precu
12	27	100.0	308	2 A36918	ARP-binding casset
13	27	100.0	311	1 JC1191	osteopontin precu
14	27	100.0	314	1 S09575	osteopontin precu
15	27	100.0	317	1 A25917	osteopontin precu
16	27	100.0	335	2 H84080	sugar ABC transpor
17	27	100.0	339	2 B75032	conserved hypotet
18	27	100.0	339	2 F96638	hypothetical prote
19	27	100.0	340	2 T26886	hypothetical prote
20	27	100.0	342	1 B71051	conserved hypotet
21	27	100.0	346	2 S77971	FWRFamide-related
22	27	100.0	350	2 B33319	-microtubule-associ
23	27	100.0	381	2 A86342	FHL16.9 protein -
24	27	100.0	383	2 T36802	probable acyl-CoA
25	27	100.0	390	2 C70765	probable cobl - My
26	27	100.0	396	1 TRYXB4	alpha-lytic protei
27	27	100.0	431	2 S45038	protein disulfide-
28	27	100.0	439	2 S19656	protein disulfide-
29	27	100.0	441	2 B84264	isochoriemate synt

30	27	100.0	445	1 QRECL	lambda receptor pr
31	27	100.0	446	2 C91256	maltose high-affin
32	27	100.0	446	2 G86096	maltose high-affin
33	27	100.0	458	2 JC5193	nuclear protein H7
34	27	100.0	469	2 JC4523	apoptosis protein
35	27	100.0	525	2 G84406	halolysin [importe
36	27	100.0	536	2 A40676	sphere erqarelle p
37	27	100.0	561	2 S35637	high mobility grou
38	27	100.0	598	2 T24318	hypothetical prote
39	27	100.0	600	2 T36195	probable acyl-CoA
40	27	100.0	648	2 C71658	probable primosoma
41	27	100.0	648	2 H97799	probable primosoma
42	27	100.0	655	2 E75551	probable acyl-pept
43	27	100.0	691	2 T45933	hypothetical prote
44	27	100.0	708	2 A41285	DNA-binding protei
45	27	100.0	709	2 T46413	hypothetical prote

## ALIGNMENTS

### RESULT 1

C83492  
hypothetical protein PA1228 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: C83492  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Sam,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: C83492  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-116 <STO>  
A:Cross-references: GB:AE004552; GB:AE0047150; NID:G9947150; PIDN:AA04617.1; GSPDB:GN00:  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA1228

Query Match 100.0%; Score 27; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|||  
Db 15 GRGDS 19

### RESULT 2

F75582  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: F75582  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: F75582  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-134 <WHI>  
A:Cross-references: GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF12529.1; PID:G646082  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRA0289  
A:Map position: 2

osteopontin precursor - chicken  
N;Alternate names: bone sialoprotein I; phosphoprotein I, secreted  
C;Species: Gallus gallus (chicken)  
C;Date: 20-Mar-1992 #sequence,revision 15-Aug-1997 #text change 21-Jul-2000  
C;Accession: I51384; A40019; A37962; S16462  
R;Rafidi, K.; Smilkina, I.; Johnson, E.; Moore, M.A.; Gerstenfeld, L.C.  
Gene 140, 163-169, 1994  
A;Title: Characterization of the chicken osteopontin-encoding gene.  
A;Reference number: I51384; MUID:94129394; PMID:8144023  
A;Accession: I51384  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-264 <R#>  
A;Cross-references: EMBL:U01844; NID:g404635; PID:AAAL8584.1; PID:g404636  
R;Castagnola, P.; Bet, P.; Quarto, R.; Genuari, M.; Cancedda, R.  
J. Biol. Chem. 266, 9944-9949, 1991  
A;Title: cDNA cloning and gene expression of chicken osteopontin. Expression of osteop  
A;Reference number: A40019, MUID:91236779; PMID:2333060  
A;Accession: A40019  
A;Molecule type: mRNA  
A;Residues: 1-215, 'P', 217-234, 'A', 236-239, 'A', 241-264 <AS>  
A;Cross-references: GB:X56772; NID:G58344; PID:CAA40091.1; PID:g4493384  
R;Moore, M.A.; Gotoh, Y.; Rafidi, K.; Gerstenfeld, L.C.  
Biochemistry 30, 2501-2508, 1991  
A;Title: Characterization of a cDNA for chicken osteopontin: expression during bone dev  
A;Reference number: A37962; MUID:91159433; PMID:2001376  
A;Accession: A37962  
A;Molecule type: mRNA  
A;Residues: 1-103, 'G', 105-264 <MO>  
A;Cross-references: GB:M59182; GB:J05323; NID:g211276; PIDN:AAAG2729.1; PID:g211277  
C;Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
C;Genetics:

C; Genomics: 18/3; 35/3; 70/3; 85/3; 164/3  
A; Introns: 18/3; 35/3; 70/3; 85/3; 164/3  
C; Superfamily: osteopontin  
C; Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
F; 1-16/Domain: signal sequence #status predicted <SIG>  
F; 17-264/Product: osteopontin #status predicted <MAT>  
F; 100-136/Region: aspartic acid-rich  
F; 138-140/Region: cell attachment (R-G-D) motif  
F; 196-109\_204\_242/Binding site: carbohydrate (Asn)  
#status predicted

C:Bronzehead, Chacko (A. J. 1987)  
C:Species: Gallus Gallus (Chicken)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Aug-1999  
C:Accession: A28512  
C:Title: Genetic analysis of the cell binding domain region of the chicken fibronectin  
R:Reference number: A28512; MUID:88050950; PMID:2823899

A:morecore type: DNA  
A:Residues: 1-273 <KUB>  
A:Cross-references: GB:X06533; NID:G63393; PIDN:CAA29781.1; PID:G295716  
A:Note: the authors translated the codon CCG for residue 190 as Gln, CAG for residue 24.  
C:Genetics:

F:150-172/Domain: fibronectin type III repeat homology <FN3>  
F:167-169/Region: cell attachment (R-G-D) motif  
F:184-266/Domain: fibronectin type III repeat homology <FN3>

Query Match 100.0%; Score 27; DB 2; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+22;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 166 GRGDS 170

RESULT 7  
 JS0638  
 Osteopontin precursor - bovine  
 N:Alternate names: bone sialoprotein I; phosphoprotein I, secreted  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Jun-1992 #sequence revision 27-Jun-1994 #text\_change 10-Sep-1999  
 A:Accession: JS0638; PC2024; PQ0160  
 R:Kerr, J.M.; Fisher, L.W.; Termini, J.D.; Young, M.F.  
 Gene 108, 237-243, 1991  
 A>Title: The cDNA cloning and RNA distribution of bovine osteopontin.  
 A:Reference number: JS0638; MUID:92084141; PMID:1721033  
 A:Accession: JS0638  
 A:Molecule type: mRNA  
 A:Residues: 1-278 <ER>  
 A:Cross-references: GB:M6236  
 A>Note: The authors translated the codon CCG for residue 221 as Ala  
 R:Sorensen, E.S.; Petersen, T.E.  
 Biochem. Biophys. Res. Commun. 198, 200-205, 1994  
 A>Title: Identification of two phosphorylation motifs in bovine osteopontin.  
 A:Reference number: JC2050; MUID:94121631; PMID:8292023  
 A:Accession: PC2024  
 A:Molecule type: protein  
 A:Residues: 17-30;36-55, 57-70;153-161;167-215;250-278 <SOR>  
 R:Prince, C.W.; Dickie, D.; Krumdieck, C.L.  
 Biochem. Biophys. Res. Commun. 177, 1205-1210, 1991  
 A>Title: Osteopontin, a substrate for transglutaminase and factor XIII activity.  
 A:Reference number: PQ0160; MUID:91282766; PMID:1676261  
 A:Accession: PQ0160  
 A:Molecule type: protein  
 A:Residues: 17-21, 'ITP' <PRI>  
 A:Experimental source: bone  
 A>Note: 22-Thr and 23-Ser were also found  
 C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
 C:Comment: This protein is thought to anchor osteoclasts to bone.  
 C:Superfamily: osteopontin  
 C:Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-278/Product: osteopontin #status experimental <NAT>  
 F:86-93/Region: aspartic acid-rich  
 F:152-154/Region: cell attachment (R-G-D) motif  
 F:23,24,26,27,60,62,63,184,188,205,210,267,272,274/Binding site: phosphate (Ser) (covalent)  
 F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:178/Binding site: phosphate (Thr) (covalent) #status experimental

Query Match 100.0%; Score 27; DB 1; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 151 GRGDS 155

RESULT 8  
 S43852  
 Osteopontin precursor - hyalomeres (polychloris penicillatus)  
 N:Contains: neurotactin Pol-RFamide I; neurotactin Pol-RFamide II  
 C:Species: Polychloris penicillatus  
 C>Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text\_change 07-May-1999  
 A:Accession: S43852; JQ2216; A30321  
 R:Schmutzler, C.; Diekhoff, D.; Grimmlikhuijzen, C.J.P.  
 Biochem. J. 299, 431-436, 1994  
 A>Title: The primary structure of the Pol-RFamide neurotactin precursor protein from th

A:Reference number: S43852; MUID:94226606; PMID:7909659  
 A:Accession: S43852  
 A:Molecule type: mRNA  
 A:Residues: 1-287 <SCH>  
 A:Cross-references: EMBL:L14777; NID:9294246; PID:9294247  
 R:Grimmelikhuijzen, C.J.P.; Rinehart, K.L.; Spencer, A.N.  
 Biochem. Biophys. Res. Commun. 183, 375-382, 1992  
 A>Title: Isolation of the neurotactin <Glu-Tyr-Leu-Lys-Gly-Arg-Phe-NH2 (Pol-RFamide II)  
 A:Reference number: JQ2216; MUID:92198411; PMID:1550547  
 A:Accession: JQ2216  
 A:Molecule type: protein  
 A:Residues: 53-59 <GRI>  
 A>Note: this sequence was confirmed by chemical synthesis  
 R:Grimmelikhuijzen, C.J.P.; Hahn, M.; Rinehart, K.L.; Spencer, A.N.  
 Brain Res. 475, 198-203, 1988  
 A>Title: Isolation of <Glu-Leu-Leu-Gly-Arg-Phe-NH2 (Pol-RFamide), a novel neurotactin  
 A:Reference number: A30321; MUID:89105774; PMID:2905621  
 A:Accession: A30321  
 A:Molecule type: protein  
 A:Residues: 182-188 <GR2>  
 A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
 A>Note: this sequence was confirmed by chemical synthesis  
 C:Keywords: amidated carboxyl end; neurotactin; pyroglutamic acid  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:53-59/Product: neurotactin Pol-RFamide II #status experimental <MI11>  
 F:55-71/Product: neurotactin Pol-RFamide II #status experimental <MI12>  
 F:78-84/Product: neurotactin Pol-RFamide II #status experimental <MI13>  
 F:91-97/Product: neurotactin Pol-RFamide II #status experimental <MI14>  
 F:104-110/Product: neurotactin Pol-RFamide II #status experimental <MI15>  
 F:117-123/Product: neurotactin Pol-RFamide II #status experimental <MI16>  
 F:130-136/Product: neurotactin Pol-RFamide II #status experimental <MI17>  
 F:143-149/Product: neurotactin Pol-RFamide II #status experimental <MI18>  
 F:169-175/Product: neurotactin Pol-RFamide II #status experimental <MI19>  
 F:182-188/Product: neurotactin Pol-RFamide I #status experimental <MA1>  
 F:195-201/Product: neurotactin #status predicted <MI11>  
 F:208-214/Product: neurotactin Pol-RFamide II #status experimental <MI1A>  
 F:221-227/Product: neurotactin Pol-RFamide II #status experimental <MI1B>  
 F:53,65,78,91,104,117,130,143,169,182,208,221/Modified site: pyrrolidone carboxylic acid  
 F:59,71,84,97,110,123,136,149,175,188,214,227/Modified site: amidated carboxyl end (Phe)

Query Match 100.0%; Score 27; DB 2; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 163 GRGDS 167

RESULT 9  
 A37818  
 Osteopontin precursor - mouse  
 N:Alternate names: bone sialoprotein I; early T lymphocyte activation 1 protein; phospho  
 C:Species: Mus musculus (house mouse)  
 C>Date: 12-Feb-1993 #sequence revision 27-Jun-1994 #text\_change 10-Sep-1999  
 A:Accession: A37818; S04078; S12064; A33853; J10105; A6031; S11677  
 R:Wiyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 J. Biol. Chem. 265, 14432-14438, 1990  
 A>Title: The mouse osteopontin gene. Expression in monocytic lineages and complete nucle  
 A:Reference number: A37818; MUID:90354433; PMID:2387863  
 A:Accession: A37818  
 A:Molecule type: DNA  
 A:Residues: 1-294 <MIY>  
 A:Cross-references: EMBL:X51834  
 R:Wiyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 Nucleic Acids Res. 17, 3298, 1989  
 A>Title: Nucleotide sequence of cDNA for mouse osteopontin-like protein.  
 A:Reference number: S04078; MUID:89263742; PMID:2726465  
 A:Accession: S04078  
 A:Molecule type: mRNA  
 A:Residues: 1-294 <MI2>  
 A:Cross-references: EMBL:X13986; NID:953755; PIDN:CNAJ2165.1; PID:953756  
 R:Yamamoto, S.

submitted to the EMBL Data Library, January 1990

A:Reference number: S12054  
 A:Accession: S12054  
 A:Molecule type: DNA  
 A:Residues: 1-121, 'F', 123-294 <YAM>  
 A:Cross-references: EMBL:X51834; NID:G53523; PIDN:CAA36132.1; PID:G297546  
 R:Craig, A.M.; Smith, J.H.; Denhardt, D.T.  
 C: Biol. Chem. 264, 9682-9689, 1989  
 A:Title: Osteopontin, a transformation-associated cell adhesion phosphoprotein, is induced by 1,25-(OH)<sub>2</sub>D<sub>3</sub> in osteoblasts  
 A:Reference number: A33853; MUID:89255479; PMID:2722855  
 A:Accession: A33853  
 A:Molecule type: mRNA  
 A:Residues: 1-98, 'G', 100-294 <CEA>  
 A:Cross-references: SB:J04806; NID:G200157; PIDN:AA57265.1; PID:G200158  
 R:Patel, R.; Freeman, G.J.; Singh, R.P.; Wei, F.Y.; Durfee, T.; Blattner, F.; Regnier, J. Exp. Med. 170, 145-161, 1989  
 A:Title: Structural and functional studies of the early T lymphocyte activation 1 (Eta-1) on.  
 A:Reference number: J10105; MUID:89310352; PMID:2783738  
 A:Accession: J10105  
 A:Molecule type: mRNA  
 A:Residues: 1-42, 'P', 44-294 <PAT>  
 A:Cross-references: EMBL:X16151; NID:G50863; PIDN:CAA34276.1; PID:G50864  
 R:Singh, R.P.; Patil, R.; Schwartz, J.; Singh, P.; Cantor, H.  
 J. Exp. Med. 171, 1933-1942, 1990  
 A:Title: Definition of a specific interaction between the early T lymphocyte activation 1 (Eta-1) and the early T lymphocyte activation 2 (Eta-2) proteins.  
 A:Reference number: A60931; MUID:90278349; PMID:2351930  
 A:Accession: A60931  
 A:Molecule type: Protein  
 A:Residues: 158-176 <SIN>  
 C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid, and serine.  
 C:Genetics: Eta-1  
 A:Gene: Eta-1  
 A:Map position: 5  
 A:Introns: 18/3; 30/3; 57/3; 71/3; 165/3  
 C:Superfamily: osteopontin  
 C:Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-294/Product: osteopontin #status predicted <SIG>  
 F:85-96/Region: aspartic acid-rich  
 F:144-146/Region: cell attachment (R-G-D) motif  
 F:78/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 27; DB 1; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5

DB 143 GRGDS 147

RESULT 10

JC5811  
 Osteopontin - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 07-May-1999  
 C:Accession: JC5811  
 R:Lasa, M.; Chang, P.J.; Prince, C.W.; Pinna, L.A.  
 Biochem. Biophys. Res. Commun. 240, 602-605, 1997  
 A:Title: Phosphorylation of osteopontin by Golgi apparatus casein kinase.  
 A:Reference number: JC5811; MUID:98063283; PMID:9398611

A:Accession: JC5811

A:Molecule type: protein

A:Residues: 1-301 <LAS>

A:Experimental source: brain

C:Comment: This protein is involved in the initiation of the bone calcification process, and early resistance to bacterial infection and binding of tumor cells at secondary sites.  
 C:Superfamily: osteopontin

Query Match 100.0%; Score 27; DB 2; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5

DB 127 GRGDS 131

RESULT 11

SEPGO

osteopontin precursor - pig  
 N:Alternate names: bone sialoprotein I; phosphoprotein I, secreted  
 N:Contains: 20K glycoprotein; 23K glycoprotein  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 10-Sep-1999  
 C:Accession: S14903; A35204; S24372; S06690  
 R:Wrana, J.L.; Zhang, Q.; Sodek, J.  
 Nucleic Acids Res. 17, 10119, 1989  
 A:Title: Full length cDNA sequence of porcine secreted phosphoprotein-I (SPP-I, osteopontin)  
 A:Reference number: S14903; MUID:90098793; PMID:2602123

A:Accession: S14903

A:Molecule type: mRNA

A:Residues: 1-303 <WRA>

A:Cross-references: EMBL:X16575; NID:G2120; PIDN:CAA34594.1; PID:G2121

R:Zhang, Q.; Domenicucci, C.; Goldberg, H.A.; Wrana, J.L.; Sodek, J.

J. Biol. Chem. 265, 7583-7589, 1990

A:Title: Characterization of fetal porcine bone sialoproteins, secreted phosphoprotein I, is derived from the carboxyl terminus of SPPI.

A:Reference number: A35204; MUID:90237064; PMID:2332443

A:Accession: A35204

A:Molecule type: protein

A:Residues: 17-36; 172-211 <ZHA>

R:Zhang, Q.; Wrana, J.L.; Sodek, J.

Eur. J. Biochem. 207, 649-659, 1992

A:Title: Characterization of the promoter region of the porcine opr. (osteopontin, secreted).

A:Reference number: S24372; MUID:92339454; PMID:1633816

A:Accession: S24372

A:Molecule type: DNA

A:Residues: 1-18 <ZHA>

A:Cross-references: EMBL:M84121; NID:G164599; PIDN:AAA31094.1; PID:G164600

C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid, and serine.

C:Superfamily: osteopontin

C:Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-303/Product: osteopontin #status predicted <SIG>

F:86-93/Region: aspartic acid-rich

F:154-156/Region: cell attachment (R-G-D) motif

F:172-303/Product: 23K glycoprotein #status experimental <23K>

F:199-303/Product: 20K glycoprotein #status experimental <20K>

F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 27; DB 1; Length 303;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5

DB 153 GRGDS 157

RESULT 12

A36918

ATP-binding cassette membrane transporter abca - Aeromonas salmonicida

C:Species: Aeromonas salmonicida

C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 02-Feb-2001

C:Accession: A36918

R:Chu, S.; Trust, T.J.

J. Bacteriol. 175, 3105-3114, 1993

A:Title: An Aeromonas salmonicida gene which influences a-protein expression in Escherichia coli.

A:Reference number: A36918; MUID:93259558; PMID:8431726

A:Accession: A36918

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <CHU>

**A**: osteopontin precursor, splice form A - human  
**N**/Alternate names: bone sialoprotein I; lactopontin; milk protein, 75K; secreted phospho-N;  
C: Species: Homo sapiens (man)  
C:/Date: 20-May-1994 #sequence revision 27-Jun-1994 #text change 16-Jun-2000  
C:/Accession: S50028; SO9575; A35326; JQ1529; I56986; I76601; I76602; A41802; S04505  
Biochim. Biophys. Acta 96, 255-262, 1994  
R/Higuchi, N.; Setoguchi, M.; Matsushima, K.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
A:/Title: Cloning and characterization of the human osteopontin gene and its promoter.  
A:/Reference number: S50028; PMID:95031968; PMID:79455249  
A:/Accession: S50028  
A:/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-314 <HID>  
A/Cross-references: EMBL:D14813; NID:g506341; PIDN:BAA03554.1; PID:g506342  
A/Note: the authors translated the codon GTG for residue 5 as Tyr  
R/Kiefer, M.C.; Bauer, D.M.; Barr, P.J.  
Nucleic Acids Res. 17, 3306, 1989  
A/Title: The cDNA and derived amino acid sequence for human osteopontin.  
A/Reference number: SO9575; PMID:89263749; PMID:2726470  
A/Accession: SO9575  
A/Molecule type: mRNA  
A/Residues: 1-314 <IE>  
A/Cross-references: EMBL:X13694; NID:g35147; PIDN:CRA31384.1; PID:g35148  
R/Yong, M.F.; Kerr, J.X.; Termini, J.D.; Wewer, U.M.; Ge Wang, M.; McBride, C.W.; Fishv Genomics 7, 491-502, 1990  
A/Title: cDNA cloning, mRNA distribution and heterogeneity, chromosomal location, and R/Biochem. Biophys. Res. Commun. 184, 859-864, 1992  
A/Title: Molecular cloning and sequencing of cDNA encoding urinary stone protein, which A/Reference number: JQ1529; PMID:92246977; PMID:I575754  
A/Accession: JQ1529  
A/Status: translation not shown  
A/Molecule type: mRNA  
A/Residues: 67-273 <KH>  
A/Experimental source: kidney  
R/Saitoh, Y.; Kuratsuji, J.; Takeshima, H.; Yamamoto, S.; Ushio, Y.  
Lab. Invest. 72, 55-63, 1995  
A/Title: Expression of Osteopontin in human glioma. Its correlation with the malignancy.  
A/Reference number: I56986; PMID:95139605; PMID:7837791  
A/Accession: I56986  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-187,'H','I',189-236,'A',238-314 <SAII>  
A/Cross-references: GB:D28759; NID:g633074; PIDN:BAA05949.1; PID:g992948  
A/Accession: I76601  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-57,72-187,'H',189-236,'A',238-314 <SAIL>  
A/Cross-references: GB:D28760; NID:g633075; PIDN:BAA05950.1; PID:g992949  
A/Accession: I76602  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-31,59-187,'H',189-236,'A',238-314 <SAIS>  
A/Cross-references: GB:D28761; NID:g633076; PIDN:BAA05951.1; PID:g992950  
R/Shiraga, H.; Min, W.; Vandusen, W.J.; Clayman, M.D.; Miner, D.; Terrell, C.H.; Sherbot Proc. Natl. Acad. Sci. U.S.A. 89, 426-430, 1992  
A/Title: Inhibition of calcium oxalate crystal growth in vitro by uropontin: another men A/Reference number: A41802; PMID:92108068; PMID:1729712  
A/Accession: A41802  
A/Molecule type: protein  
A/Residues: 19-62 <SHI>  
R/Senger, D.R.; Perruzzi, C.A.; Papadopoulos, A.; Tenen, D.G.  
Biochim. Biophys. Acta 996, 43-48, 1989  
A/Title: Purification of a human milk protein closely similar to tumor-secreted phosphor A/Reference number: S04505; PMID:89287357; PMID:2736258  
A/Accession: S04505  
A/Status: preliminary

A:Molecule type: protein  
A:Residues: 17-23;169-182 <SEN>  
C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
C:Genetics:  
A:Gene: GDB:SPPI; BNSP: OPN  
A:Cross-references: GDB:118889; OMIM:166490  
A:Map position: 4q21-q25  
C:Superfamily: osteopontin  
C:Keywords: alternative splicing; bone; cell binding; extracellular matrix; phosphoprotein  
F:1-314/Product: osteopontin precursor, splice form A #status predicted <PRA>  
F:1-57,72-314/Product: osteopontin precursor, splice form B #status predicted <PRB>  
F:1-31,59-314/Product: osteopontin precursor, splice form C #status predicted <PRC>  
F:1-16/Domains: signal sequence #status predicted <SIG>  
F:17-314/Product: osteopontin #status predicted <MAT>  
F:86-95/Region: aspartic acid-rich  
F:159-163/Region: cell attachment (R-G-D) motif  
F:24,26,27,62,63,191,195,234,303,308,310/Binding site: phosphate (Ser) (covalent) #status  
F:79,106/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:185/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 100.0%; Score 27; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 5;

QY 1 GRGDS 5  
Db 158 GRGDS 162

RESULT 15  
A25917  
osteopontin precursor - rat  
N:Alternate names: bone sialoprotein I; phosphoprotein I, secreted; tumor-secreted phosph  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 18-Dec-1987 #sequence revision 27-Jun-1994 #text change 10-Sep-1999  
C:Accession: A25917; A45132; B45132; S28772; S04506; A45925  
R:Oldberg, A.; Franzen, A.; Heinegard, D.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8919-8823, 1986  
A:Title: Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveal  
A:Reference number: A25917; MUID:87067405; PMID:3024151  
A:Accession: A25917  
A:Molecule type: mRNA  
A:Residues: 1-317 <OLD>  
A:Cross-references: GB:X14656; NID:G205859; PIDN:AAA1762.1; PID:G205860  
R:Singh, K.; Mukherjee, A.B.; De Vouge, M.W.; Mukherjee, B.B.  
J. Biol. Chem. 267, 23847-23851, 1992  
A:Title: Differential processing of osteopontin transcripts in rat kidney- and osteoblas  
A:Reference number: A45132; MUID:93054745; PMID:1429723  
A:Accession: A45132  
A:Molecule type: protein  
A:Residues: 36-51 <SIM1>  
A:Experimental source: kidney  
A>Note: sequence extracted from NCBI backbone (NCBIP:118869)  
A:Accession: B45132  
A:Molecule type: protein  
A:Residues: 272-282 <SIM2>  
A>Note: sequence extracted from NCBI backbone (NCBIP:118871)  
R:Prince, C.W.; Oosawa, T.; Butler, W.T.; Tomana, M.; Brown, A.S.; Brown, M.; Schrohenlo  
J. Biol. Chem. 262, 2900-2907, 1987  
A:Title: Isolation, characterization, and biosynthesis of a phosphorylated glycoprotein  
A:Reference number: S28772; MUID:87137549; PMID:3469201  
A:Accession: S28772  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 17-26, X' <PRI>  
R:Senger, D.R.; Perruzzi, C.A.; Papadopoulos, A.; Teren, D.G.  
Biochim. Biophys. Acta 996, 43-48, 1989  
A:Title: Purification of a human milk protein closely similar to tumor-secreted phosphop  
A:Reference number: S04505; MUID:89287357; PMID:2736258  
A:Accession: S04506  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 17-26;155-167 <SEN>

R:Senger, D.R.; Perruzzi, C.A.; Gracey, C.F.; Papadopoulos, A.; Teren, D.G.  
Cancer Res. 48, 5770-5774, 1988  
A:Title: Secreted phosphoproteins associated with neoplastic transformation: close homo  
A:Reference number: A45925; MUID:89002730; PMID:3167835  
A:Accession: A45925  
A:Molecule type: protein  
A:Residues: 17-25 <SE2>  
C:Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
C:Superfamily: osteopontin  
C:Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
F:1-16/Domains: signal sequence #status predicted <SIG>  
F:17-317/Product: osteopontin #status predicted <MAT>  
F:86-96/Region: aspartic acid-rich  
F:144-146/Region: cell attachment (R-G-D) motif  
F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 27; DB 1; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
Db 143 GRGDS 147

Search completed: April 16, 2004, 07:29:19  
Job time : 12.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Ran on: April 16, 2004, 07:28:08 ; Search time 33.625 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588B-4  
Perfect score: 27  
Sequence: 1 GRGDS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PTCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	5	9	US-09-096-749A-72
2	27	100.0	5	10	US-09-903-412-72
3	27	100.0	5	11	US-09-991-588B-4
4	27	100.0	5	12	US-10-169-085-22
5	27	100.0	5	12	US-09-961-381A-2
6	27	100.0	5	13	US-10-154-507-4
7	27	100.0	5	14	US-10-303-583-5
8	27	100.0	5	14	US-10-174-717A-72
9	27	100.0	5	14	US-10-165-155-72
10	27	100.0	5	14	US-10-190-162-72
11	27	100.0	5	14	US-10-042-696-34
12	27	100.0	5	16	US-10-208-894A-4
13	27	100.0	6	9	US-09-892-071-6
14	27	100.0	6	9	US-09-837-969A-28
15	27	100.0	6	9	US-09-837-969A-46

16	27	100.0	6	9	US-09-888-260-15
17	27	100.0	6	9	US-09-961-834-1
18	27	100.0	6	9	US-09-364-597A-21
19	27	100.0	6	9	US-09-177-843-1
20	27	100.0	6	9	US-09-841-321A-28
21	27	100.0	6	9	US-09-841-321A-46
22	27	100.0	6	9	US-09-871-374-3
23	27	100.0	6	9	US-09-972-772-31
24	27	100.0	6	11	US-09-991-588B-7
25	27	100.0	6	12	US-09-961-381A-3
26	27	100.0	6	12	US-09-380-682-63
27	27	100.0	6	12	US-10-136-187-42
28	27	100.0	6	13	US-10-001-345-31
29	27	100.0	6	13	US-10-007-270-32
30	27	100.0	6	13	US-10-007-270-35
31	27	100.0	6	13	US-10-114-176-2
32	27	100.0	6	14	US-10-046-801-5
33	27	100.0	6	14	US-10-237-850-83
34	27	100.0	6	14	US-10-138-335-31
35	27	100.0	6	14	US-10-262-435-5
36	27	100.0	6	14	US-10-325-021-9
37	27	100.0	6	14	US-10-323-013-2
38	27	100.0	6	14	US-10-439-532-15
39	27	100.0	6	14	US-10-086-208-5
40	27	100.0	6	15	US-10-259-609-14
41	27	100.0	6	15	US-10-162-432-31
42	27	100.0	7	8	US-08-987-756-1
43	27	100.0	7	9	US-09-364-597A-22
44	27	100.0	7	10	US-09-911-569-8
45	27	100.0	7	11	US-09-991-588B-10

## ALIGNMENTS

RESULT 1  
US-09-096-749A-72  
; Sequence 72, Application US/09096749A  
; Patent No. US20020019517A1  
; GENERAL INFORMATION:  
; APPLICANT: Koieda, Shohei  
; TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES  
; NUMBER OF SEQUENCES: 118  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth P.A.  
; STREET: 121 South Eighth Street, Ste. 1600  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0b  
; CURRENT APPLICATION DATA: US/09/096,749A  
; APPLICATION NUMBER: US/09/096,749A  
; FILING DATE: June 12, 1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ann S. Viksnias  
; REGISTRATION NUMBER: 37,748  
; REFERENCE/DOCKET NUMBER: 109.034US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 373-6900  
; TELEFAX: (612) 339-3061  
; INFORMATION FOR SEQ ID NO: 72:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

;  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-09-096-749A-72

Query Match 100.0%; Score 27; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5

## RESULT 2

US-09-903-412-72  
; Sequence 72, Application US/09903412  
; Publication No. US20030027319A1  
; GENERAL INFORMATION:  
; APPLICANT: Koide, Shohei  
; TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES  
; FILE REFERENCE: 129,050US1  
; CURRENT APPLICATION NUMBER: US/09/903,412  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 60/217,474  
; PRIOR FILING DATE: 2000-07-11  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 72  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: The sequence of the FG loop of WT from library #4.  
US-09-903-412-72

Query Match 100.0%; Score 27; DB 10; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5

## RESULT 3

US-09-991-588B-4  
; Sequence 4, Application US/09991588B  
; Publication No. US20030219429A1  
; GENERAL INFORMATION:  
; APPLICANT: Eudny, John A.  
; TITLE OF INVENTION: Composition and Method for Bone Regeneration  
; FILE REFERENCE: 1008-120 US  
; CURRENT APPLICATION NUMBER: US/09/991,588B  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 09/122,348  
; PRIOR FILING DATE: 1998-07-24  
; NUMBER OF SEQ ID NOS: 23  
; SEQ ID NO 4  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Purchased commercially or sequence is synthesized  
US-09-991-588B-4

Query Match 100.0%; Score 27; DB 11; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5

## RESULT 4

US-10-169-085-22  
; Sequence 22, Application US/10169085  
; Publication No. US20040033215A1  
; GENERAL INFORMATION:  
; APPLICANT: Kobayashi, Nacya  
; APPLICANT: Philippe, Leboulch  
; APPLICANT: Tanaka, No. US20040033215A1  
; APPLICANT: Fujiwara, Toshiyoshi  
; APPLICANT: Toshihori, Totusugawa  
; TITLE OF INVENTION: METHOD FOR PROLIFERATING A LIVER CELL, A LIVER CELL OBTAINED  
; TITLE OF INVENTION: THEREBY, AND USE THEREOF  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/169,085  
; CURRENT FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: US/09/809,187  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 23  
; SEQ ID NO 22  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Cell adhesion peptide  
US-10-169-085-22

Query Match 100.0%; Score 27; DB 12; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5

## RESULT 5

US-09-961-381A-2  
; Sequence 2, Application US/09961381A  
; Publication No. US20020061515A1  
; GENERAL INFORMATION:  
; APPLICANT: Lynch, Gary  
; APPLICANT: Bi, Xiaoning  
; APPLICANT: Gall, Christine M.  
; TITLE OF INVENTION: Model for Neurodegenerative Diseases Involving Amyloid  
; TITLE OF INVENTION: Accumulation  
; FILE REFERENCE: 1819,0040001  
; CURRENT APPLICATION NUMBER: US/09/961,381A  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/235,374  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-09-961-381A-2

Query Match 100.0%; Score 27; DB 12; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
|  
|  
|  
|  
Db 1 GRGDS 5



## RESULT 6

US-10-154-507-4  
 ; Sequence 4, Application US/10154507  
 ; Publication No. US20020147138A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Firestone, Raymond A.  
 ; TITLE OF INVENTION: Enzyme-Activated Anti-Tumor Prodrug Compounds  
 ; FILE REFERENCE: 9/145-1-D1  
 ; CURRENT APPLICATION NUMBER: US/10/154,507  
 ; CURRENT FILING DATE: 2002-05-23  
 ; PRIOR APPLICATION NUMBER: US 09/551,147  
 ; PRIOR FILING DATE: 2000-04-15  
 ; PRIOR APPLICATION NUMBER: US 60/134,135  
 ; PRIOR FILING DATE: 1999-05-14  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)..(5)  
 ; OTHER INFORMATION: Adhesion peptide  
 US-10-154-507-4

Query Match 100.0%; Score 27; DB 13; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 7

US-10-303-583-5  
 ; Sequence 5, Application US/10303583  
 ; Publication No. US20030099685A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkar, Sammy  
 ; TITLE OF INVENTION: Osteopontin Coated Surfaces and Methods of Use  
 ; FILE REFERENCE: cme-100CP  
 ; CURRENT APPLICATION NUMBER: US/10/303,583  
 ; CURRENT FILING DATE: 2002-11-22  
 ; PRIOR APPLICATION NUMBER: US/09/134,253  
 ; PRIOR FILING DATE: 1998-08-14  
 ; PRIOR APPLICATION NUMBER: 08/916,912  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-303-583-5

Query Match 100.0%; Score 27; DB 14; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 8

US-10-174-717A-72  
 ; Sequence 72, Application US/10174717A  
 ; Publication No. US20030108948A1  
 ; APPLICANT: Koide, Shohei  
 ; TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES

NUMBER OF SEQUENCES: 118  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth P.A.  
 STREET: 121 South Eighth Street, Ste. 1600  
 CITY: Minneapolis  
 STATE: MN  
 COUNTRY: USA  
 ZIP: 55402

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: WINDOWS  
 SOFTWARE: FastSeq Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/174,717A  
 FILING DATE: 18-Jun-2002  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/096,749  
 FILING DATE: June 12, 1998  
 APPLICATION NUMBER: 60/049,410  
 FILING DATE: June 12, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ann S. Viksnins  
 REGISTRATION NUMBER: 37,748  
 REFERENCE/DOCKET NUMBER: 109.034US4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (612) 373-6900  
 TELEFAX: (612) 339-3061

INFORMATION FOR SEQ ID NO: 72:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 ORIGINAL SOURCE:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 72:  
 US-10-174-717A-72

Query Match 100.0%; Score 27; DB 14; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
 Db 1 GRGDS 5

## RESULT 9

US-10-165-155-72  
 ; Sequence 72, Application US/10165155  
 ; Publication No. US20030134386A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koieda, Shohei

TITLE OF INVENTION: ARTIFICIAL ANTIBODY POLYPEPTIDES  
 NUMBER OF SEQUENCES: 118  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth P.A.  
 STREET: 121 South Eighth Street, Ste. 1600  
 CITY: Minneapolis  
 STATE: MN  
 COUNTRY: USA  
 ZIP: 55402

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/165,155

```
;
;
; PILING DATE: 06-Jun-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,749
; FILING DATE: June 12, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ann S. Viksnins
; REGISTRATION NUMBER: 37,748
; REFERENCE/DOCKET NUMBER: 109.034US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (612) 373-6900
; TELEFAX: (612) 339-3061
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-165-155-72

Query Match 100.0%; Score 27; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 1 GRGDS 5

RESULT 11
US-10-042-696-34
; Sequence 34, Application US/10042696
; Publication No. US20030171298A1
; GENERAL INFORMATION:
; APPLICANT: Tuszynski, George
; APPLICANT: Williams, Taffy
; APPLICANT: Actor, Paul
; TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT THROMBOSPONDIN A
; FILE REFERENCE: 07206-0021
; CURRENT APPLICATION NUMBER: US/10/042,696
; CURRENT FILING DATE: 2001-10-03
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver 2.0
; SEQ ID NO 34
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic fragment/analogs of
; OTHER INFORMATION: thrombospondin
US-10-042-696-34

Query Match 100.0%; Score 27; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 1 GRGDS 5

RESULT 12
US-10-208-894A-4
; Sequence 4, Application US/10208894A
; Publication No. US20040023391A1
; GENERAL INFORMATION:
; APPLICANT: FANG, YE
; APPLICANT: LAI, FANG
; APPLICANT: PICARD, LAURENT A.G.
; APPLICANT: WEBB, BRIAN L.
; TITLE OF INVENTION: METHOD AND DEVICE FOR PROTEIN DELIVERY INTO CELLS
; FILE REFERENCE: SP02-169
; CURRENT APPLICATION NUMBER: US/10/208,894A
; CURRENT FILING DATE: 2002-07-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic RGD
; OTHER INFORMATION: peptide
US-10-208-894A-4
```

Query Match 100.0%; Score 27; DB 16; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
| | | | |  
Db 1 GRGDS 5

## RESULT 13

US-09-892-071-6  
; Sequence 6, Application US/09892071  
; Patent No. US20020028913A1  
; GENERAL INFORMATION:  
; APPLICANT: Pierschbacher, Michael D.  
; APPLICANT: Ruoslahti, Erkki I.  
; TITLE OF INVENTION: Conformationally Stabilized Cell  
; FILE REFERENCE: P-LA 3637  
; CURRENT APPLICATION NUMBER: US/09/892,071  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: US 09/366,991  
; PRIOR FILING DATE: 1999-08-04  
; PRIOR APPLICATION NUMBER: US 08/124,992  
; PRIOR FILING DATE: 1993-09-21  
; PRIOR APPLICATION NUMBER: US 08/048,576  
; PRIOR FILING DATE: 1993-04-15  
; PRIOR APPLICATION NUMBER: US 07/803,797  
; PRIOR FILING DATE: 1991-12-06  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide

## US-09-892-071-6

Query Match 100.0%; Score 27; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
| | | | |  
Db 1 GRGDS 5

## RESULT 14

US-09-837-969A-28  
; Sequence 28, Application US/09837969A  
; Patent No. US20020038150A1  
; GENERAL INFORMATION:  
; APPLICANT: Urry, Dan  
; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
; FILE REFERENCE: BERL-020/03US  
; CURRENT APPLICATION NUMBER: US/09/837,969A  
; CURRENT FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 09/258,723  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: US 60/087155  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: US 60/076297  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: {1}..{6}

; OTHER INFORMATION: Synthetic  
US-09-837-969A-28

Query Match 100.0%; Score 27; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
| | | | |  
Db 1 GRGDS 5

## RESULT 15

US-09-837-969A-46  
; Sequence 46, Application US/09837969A  
; Patent No. US20020038150A1  
; GENERAL INFORMATION:  
; APPLICANT: Urry, Dan  
; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
; FILE REFERENCE: BERL-020/03US  
; CURRENT APPLICATION NUMBER: US/09/837,969A  
; CURRENT FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 09/258,723  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: US 60/087155  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: US 60/076297  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 46  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: {1}..{6}  
; OTHER INFORMATION: Synthetic

## US-09-837-969A-46

Query Match 100.0%; Score 27; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
| | | | |  
Db 1 GRGDS 5

Search completed: April 16, 2004, 08:03:16  
Job time : 33.625 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time 8.75 Seconds  
(without alignments)  
29.754 Million cell updates/sec

Title: us-09-991-588b-4

Perfect score: 27

Sequence: 1 GRGDS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	264	1	OSTP_CHICK
2	27	100.0	278	1	OSTP_BOVIN
3	27	100.0	278	1	OSTP_SHEEP
4	27	100.0	281	1	HIS1_CORGL
5	27	100.0	287	1	PRFA_POLPE
6	27	100.0	294	1	OSTP_MOUSE
7	27	100.0	303	1	OSTP_PIG
8	27	100.0	308	1	ABCA_AERSA
9	27	100.0	311	1	OSTP_RABIT
10	27	100.0	314	1	OSTP_HUMAN
11	27	100.0	317	1	OSTP_RAT
12	27	100.0	346	1	PAR2_HELAS
13	27	100.0	350	1	ENP2_TORCA
14	27	100.0	390	1	CHLA_MYCTU
15	27	100.0	397	1	PRLA_LYSEN
16	27	100.0	431	1	PDA6_RAT
17	27	100.0	439	1	PDA6_MESAU
18	27	100.0	446	1	LAMB_ECO57
19	27	100.0	446	1	LAMB_ECOL6
20	27	100.0	446	1	LAMB_ECOL1
21	27	100.0	485	1	MKR4_HUMAN
22	27	100.0	536	1	SPH1_XENLA
23	27	100.0	561	1	SRP1_RAT
24	27	100.0	648	1	PRIA_RICPR
25	27	100.0	692	1	NEK8_HUMAN
26	27	100.0	708	1	SRP_MOUSE
27	27	100.0	721	1	ENP1_TORCA
28	27	100.0	914	1	ORC1_YEAST
29	27	100.0	1011	1	PERQ_MOUSE
30	27	100.0	1017	1	KCH4_HUMAN
31	27	100.0	1017	1	TRL3_HUMAN
32	27	100.0	1021	1	MAPA_MOUSE
33	27	100.0	1167	1	ZO2_MOUSE

34	27	100.0	1174	1	ZO2_CANFA
35	27	100.0	1190	1	ZO2_HUMAN
36	27	100.0	1245	1	POLS_SINDO
37	27	100.0	1245	1	POLS_SINDV
38	27	100.0	1256	1	FINC_CHICK
39	27	100.0	1328	1	FINC_PLEWA
40	27	100.0	1353	1	CYA9_HUMAN
41	27	100.0	2265	1	FINC_BOVIN
42	27	100.0	2386	1	FINC_HUMAN
43	27	100.0	2459	1	MAPB_RAT
44	27	100.0	2464	1	MAPB_MOUSE
45	27	100.0	2468	1	MAPB_HUMAN

## ALIGNMENTS

### RESULT 1

ID	OSTP_CHICK	STANDARD	PRT	264 AA.
AC	P23498			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Osteopontin precursor (Bone sialoprotein 1).			
GN	SPPI.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91236779; PubMed=2033080;			
RA	Castagnola P., Set P., Quarto R., Gemari M., Migliaccio G.,			
RA	Cancedda R.;			
RT	"cDNA cloning and gene expression of chicken osteopontin. Expression			
RT	of osteopontin mRNA in chondrocytes is enhanced by trypsin treatment			
RT	of cells.";			
RL	J. Biol. Chem. 266:9944-9949 (1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91159433; PubMed=2001376;			
RA	Moore M.A., Gotoh Y., Rafidi K., Gerstenfeld L.C.;			
RT	"Characterization of a cDNA for chicken osteopontin: expression			
RT	during bone development, osteoblast differentiation, and tissue			
RT	distribution.";			
RL	Biochemistry 30:2501-2508 (1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94192994; PubMed=8144023;			
RA	Rafidi K., Simikina I., Johnson E., Moore M.A., Gerstenfeld L.C.;			
RT	"Characterization of the chicken osteopontin-encoding gene.";			
RL	Gene 140:163-169 (1994).			
CC	!- FUNCTION: Bands tightly to hydroxyapatite. Appears to form an			
CC	integral part of the mineralized matrix. Probably important to			
CC	cell-matrix interaction.			
CC	!- FUNCTION: Acts as a cytokine involved in enhancing production of			
CC	interferon-gamma and interleukin-12 and reducing production of			
CC	interleukin-10 and is essential in the pathway that leads to type			
CC	I immunity (by similarity).			
CC	!- SUBUNIT: Ligand for integrin alpha-V/beta-3.			
CC	!- SUBCELLULAR LOCATION: Secreted.			
CC	!- PTM: Extensively phosphorylated on serine residues.			
CC	!- PTM: N- and O-glycosylated.			
CC	!- SIMILARITY: Belongs to the osteopontin family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

Fri Apr 16 11:04:13 2004

us-09-991-588b-4.rsp

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56772; CAA40091.1; --
CC DE EMBL; M59182; AAA62729.1; --
CC DR EMBL; U01844; AAL14584.1; --
CC DR PIR; I51384; A40019.
CC DR InterPro; IPR002038; Osteopontin.
CC DR Pfam; PF00865; Osteopontin; 1.
CC DR SMART; SM00017; OSTEO; 1.
CC DR PROSITE; PS00884; OSTEOPONTIN; 1.
CC DR Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;
KW Phosphorylation; Signal.
FT SIGNAL 1 16
FT CHAIN 17 264
FT SITE 132 134
FT CARBOHYD 106 106
FT CARBOHYD 109 109
FT CARBOHYD 204 204
FT CARBOHYD 242 242
FT CONFLICT 104 104
FT CONFLICT 216 216
FT CONFLICT 235 235
FT CONFLICT 240 240
FT CONFLICT 240 240
FT SEQUENCE 264 AA; 29142 MW; CA2F17DB70A92591 CRC64;
FT SIGNAL 1 16
FT CHAIN 17 264
FT SITE 132 134
FT CARBOHYD 106 106
FT CARBOHYD 109 109
FT CARBOHYD 204 204
FT CARBOHYD 242 242
FT CONFLICT 104 104
FT CONFLICT 216 216
FT CONFLICT 235 235
FT CONFLICT 240 240
FT CONFLICT 240 240
FT SEQUENCE 264 AA; 29142 MW; CA2F17DB70A92591 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 137 GRGDS 141

RESULT 2
OSTP_BOVIN STANDARD; PRT; 278 AA.
AC P31096; Q86PS6;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Osteopontin precursor (Bone sialoprotein 1).
GN SPPI OR OPN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92084141; PubMed=721033;
RA Kerr J.M., Fisher L.W., Termino J.D., Young M.F.;
RT "The cDNA cloning and RNA distribution of bovine osteopontin.";
RL Gene 108:237-243(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Korean; TISSUE=Kidney;
RA Lee T.Y., Ju S.K., Nam M.S.;
RT "Cloning of osteopontin (OPN) in Korean native cattle.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 17-22; 44-54; 168-182 AND 221-243.
RX TISSUE=Milk;
RA MEDLINE=93308294; PubMed=8320368;
RA Soerensen E.S., Petersen T.E.;
RT "Purification and characterization of three proteins isolated from
the protease peptone fraction of bovine milk.";
RL J. Dairy Res. 60:189-197(1993).
RN [4]
RP POST-TRANSLATIONAL MODIFICATIONS.
RX TISSUE=Milk;
RA MEDLINE=96117654; PubMed=8535240;
RA Soerensen E.S., Hoejrup P., Petersen T.E.;

```

```

RT *Posttranslational modifications of bovine osteopontin:
RT identification of twenty-eight phosphorylation and three
RT O-glycosylation sites."
RL Protein Sci. 4:2040-2049(1995).
CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an
CC integral part of the mineralized matrix. Probably important to
CC cell-matrix interaction.
CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of
CC interferon-gamma and interleukin-12 and reducing production of
CC interleukin-10 and is essential in the pathway that leads to type
CC I immunity (by similarity)
CC -!- SUBUNIT: Ligand for integrin alpha-w/beta-3.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: BONE, INNER EAR, KIDNEY, UTERUS, LUNG, BRAIN,
CC EPIDERMIS.
CC -!- SIMILARITY: Belongs to the osteopontin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M66236; AAA30462.1; --
CC DE EMBL; AF492837; AAL99081.1; --
CC GO; GO:0005576; C:extracellular; ISS.
CC GO; GO:0005125; F:cytokine activity; ISS.
CC InterPro; IPR002038; Osteopontin.
CC Pfam; PF00865; Osteopontin; 1.
CC PRINTS; PR00216; OSTEOPONTIN.
CC SMART; SM00017; OSTEO; 1.
CC PROSITE; PS00884; OSTEOPONTIN; 1.
CC Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;
KW Phosphorylation; Signal.
FT SIGNAL 1 16
FT CHAIN 17 278
FT SITE 152 154
FT MOD_RES 23 23
FT MOD_RES 24 24
FT MOD_RES 26 26
FT MOD_RES 27 27
FT MOD_RES 60 60
FT MOD_RES 62 62
FT MOD_RES 63 63
FT MOD_RES 76 76
FT MOD_RES 78 78
FT MOD_RES 81 81
FT MOD_RES 95 95
FT MOD_RES 100 100
FT MOD_RES 103 103
FT MOD_RES 115 115
FT MOD_RES 121 121
FT MOD_RES 124 124
FT MOD_RES 178 178
FT MOD_RES 184 184
FT MOD_RES 188 188
FT MOD_RES 205 205
FT MOD_RES 210 210
FT MOD_RES 233 233
FT MOD_RES 240 240
FT MOD_RES 245 245
FT MOD_RES 256 256
FT MOD_RES 267 267
FT MOD_RES 272 272
FT MOD_RES 274 274
FT CARBOHYD 131 131
FT CARBOHYD 140 140
FT CARBOHYD 145 145
FT CONFLICT 42 42
FT CONFLICT 56 56
FT CONFLICT 278 AA; EAB6EA39A6DD8E6F CRC64;
SQ SEQUENCE 278 AA; 33946 MW; EAB6EA39A6DD8E6F CRC64;

```

Query Match 100.0%; Score 27; DB 1; Length 278;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
DB 151 GRGDS 155

RESULT 3  
OSTP SHEEP STANDARD; PRT; 278 AA.

AC QXSY9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Osteopontin precursor (Bone sialoprotein 1).  
GN SP1 OR OPN.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endometrium;  
RX MEDLINE=99423513; PubMed=10491620;  
RA Johnson G.A., Spencer T.E., Burghardt R.C., Bazer F.W.;  
RT "Ovine osteopontin: I. Cloning and expression of messenger ribonucleic acid in the uterus during the perimplantation period.";  
RL Biol. Reprod. 61:884-891 (1999).  
CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an integral part of the mineralized matrix. Probably important to cell-matrix interaction.  
CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of interferon-gamma and interleukin-12 and reducing production of interleukin-10 and is essential in the pathway that leads to type I immunity (By similarity).  
CC -!- SUBUNIT: Ligand for integrin alpha-V/beta-3.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: Extensively phosphorylated on serine residues (By similarity).  
CC -!- PTM: N- and O-glycosylated (By similarity).  
CC -!- SIMILARITY: Belongs to the osteopontin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF152416; AAD3388.1; -  
GO; GO:0005576; C:extracellular; ISS.  
GO; GO:0005125; F:Cytokine activity; ISS.  
InterPro; IPR002038; Osteopontin.  
Pfam; PF00865; Osteopontin; 1.  
PRINTS; PR02216; OSTOPONTIN.  
SMART; SM00317; OSTEO; 1.  
PROSITE; PS00884; OSTROPONTIN; 1.  
KW Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion; Phosphorylation; Signal.  
FT SIGNAL 1 16 BY SIMILARITY.  
FT CHAIN 17 278 OSTROPONTIN.  
FT SITE 152 154 CELL ATTACHMENT SITE.  
FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 278 AA; 31052 MW; 37D49E1DD1FBD47 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 278;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
DB 151 GRGDS 155

RESULT 4  
HIS1\_CORGL STANDARD; PRT; 281 AA.

AC Q92472;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP phosphoribosyltransferase (EC 2.4.2.17).  
GN HISG OR CGL1504.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13059 / AS019;  
RA Kwon J.H., Jung S.I., Lee M.-S.;  
RT "Molecular cloning of the hisG gene from Corynebacterium glutamicum.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: 1-(5-phospho-D-riboseyl)-ATP + diphosphate = ATP + 5-phospho-alpha-D-ribose 1-diphosphate.  
CC -!- PATHWAY: Histidine biosynthesis; first step. Very important in the regulation of histidine metabolism.  
CC -!- SUBUNIT: Homohexamer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family. Long subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF050166; AAD02497.1; -  
EMBL; AP005278; BAB98997.1; -  
HMAP; MF 00079; - 1  
InterPro; IPR001348; ATP\_phospho\_trans.  
Pfam; PF01634; HisG; 1.  
ProDom; PD003516; ATP\_phospho\_trans; 1.  
TIGRFAMs; TIGR00070; hisG; 1.  
DR PROSITE; PS01316; ATP\_P\_PHORIBOSYLTR; 1.  
DR PRCSITE; PS01316; ATP\_P\_PHORIBOSYLTR; 1.  
KW Histidine biosynthesis; Transferase; Glycosyltransferase;  
Complete proteome.  
FT CONFLICT 10 10 G -> A (IN REF. 1).  
FT CONFLICT 24 24 G -> A (IN REF. 1).  
FT CONFLICT 74 74 MISSING (IN REF. 1).  
FT CONFLICT 99 99 D -> Y (IN REF. 1).  
FT CONFLICT 113 113 A -> R (IN REF. 1).  
FT CONFLICT 163 165 RQQ -> LT (IN REF. 1).  
FT CONFLICT 197 197 L -> C (IN REF. 1).  
SQ SEQUENCE 281 AA; 30229 MW; 62DF8724CBE9B78C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
DB 151 GRGDS 155







RX MEDLINE=90098793; PubMed=2602123;  
RA Wana J.L., Zhang Q., Sodek J.;  
RT "Full length cDNA sequence of porcine secreted phosphoprotein-I  
(SPP-I, osteopontin).";  
RL Nucleic Acids Res. 17:10119-10119 (1989).  
RN [12]  
RP SEQUENCE OF 1-18 FROM N.A.  
RX MEDLINE=92339454; PubMed=1633846;  
RA Zhang Q., Wana J.L., Sodek J.;  
RT "Characterization of the promoter region of the porcine opn  
(osteopontin, secreted phosphoprotein 1) gene. Identification of  
positive and negative regulatory elements and a 'silent' second  
promoter.";  
RL Eur. J. Biochem. 207:649-659 (1992).  
RN [3]  
RP SEQUENCE OF 17-36 AND 172-211.  
RC TISSUE=Bone;  
RX MEDLINE=90237064; PubMed=2332443;  
RA Zhang Q., Domenicucci C., Goldberg H.A., Wana J.L., Sodek J.;  
RT "Characterization of fetal porcine bone sialoproteins, secreted  
phosphoprotein I (SPP-I, osteopontin), bone sialoprotein, and a 23-kDa  
glycoprotein. Demonstration that the 23-kDa glycoprotein is derived  
from the carboxyl terminus of SPP-I";  
RL J. Biol. Chem. 265:7593-7598 (1990).  
CC -1- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an  
integral part of the mineralized matrix. Probably important to  
cell-matrix interaction.  
CC -1- FUNCTION: Acts as a cytokine involved in enhancing production of  
interferon-gamma and interleukin-12 and reducing production of  
interleukin-10 and is essential in the pathway that leads to type  
I immunity (By similarity).  
CC -1- SUBUNIT: Ligand for integrin alpha-v/beta-3.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- PTM: Extensively phosphorylated on serine residues.  
CC -1- PTM: N- and O-glycosylated.  
CC -1- SIMILARITY: Belongs to the osteopontin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
EMBL; X16575; CRA34594.1; -  
EMBL; M8412; AAA31094.1; -  
PIR; S14903; GEPGO.  
DR GO; GO:0005578; C:extracellular matrix; ISS.  
DR GO; GO:0008189; E:apoptosis inhibitor activity; ISS.  
DR GO; GO:0042056; F:chemoattractant activity; ISS.  
DR GO; GO:0005125; F:cytokine activity; ISS.  
DR GO; GO:0003793; F:defense/immunity protein activity; ISS.  
DR GO; GO:0008083; F:growth factor activity; ISS.  
DR GO; GO:0005178; F:integrin binding; ISS.  
DR GO; GO:0006916; F:anti-apoptosis; ISS.  
DR GO; GO:0007267; P:cell-cell signaling; ISS.  
DR GO; GO:0007160; P:cell-matrix adhesion; ISS.  
DR GO; GO:0030595; P:immune cell chemotaxis; ISS.  
DR GO; GO:0006954; P:inflammatory response; ISS.  
DR GO; GO:0030502; P:negative regulation of bone mineralization; ISS.  
DR GO; GO:0042102; P:positive regulation of T-cell proliferation; ISS.  
DR GO; GO:0045637; P:regulation of myeloid blood cell differenti. . .; ISS.  
DR GO; GO:0042088; P:T-helper 1 type immune response; ISS.  
DR InterPro; IPR002038; Osteopontin.  
DR Pfam; PF00865; Osteopontin; 1.  
DR PRINTS; PR00216; OSTEOPONTIN.  
DR SMART; SM00017; OSTEO; 1.  
DR PROSITE; PS00884; OSTEOPONTIN; 1.  
DR Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;  
KW Phosphorylation; Signal.  
FT SIGNAL 1 16 POTENTIAL.  
FT CHAIN 17 303 OSTEOPONTIN.  
  
FT SITE 154 156 CELL ATTACHMENT SITE.  
FT CARBOHYD 79 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 303 AA; 33668 MW; 87D6240E24013BF7 CRC64;  
  
Query Match 100.0%; Score 27; DB 1; Length 303;  
Best Local Similarity 100.0%; Pred. No. 54; Indels 0; Gaps 0;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDS 5  
|||||  
DB 153 GRGDS 157  
  
RESULT 8  
ABCA AERSA STANDARD; PRT; 308 AA.  
ID ABCA AERSA STANDARD; PRT; 308 AA.  
AC Q07698; 1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ABC transporter protein abca.  
GN ABCA.  
OS Aeromonas salmonicida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;  
OC Aeromonadaceae; Aeromonas.  
OC NCBI\_TaxID=645;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A450;  
RX MEDLINE=93259598; PubMed=8491726;  
RA Chu S., Trust T.J.;  
RT "An Aeromonas salmonicida gene which influences a protein expression  
in Escherichia coli encodes a protein containing an ATP-binding  
cassette and maps beside the surface array protein gene.";  
RL J. Bacteriol. 175:3105-3114 (1993).  
CC -1- FUNCTION: INFLUENCES THE EXPRESSION OF THE SURFACE ARRAY PROTEIN  
GENE (VAPA). MAY HAVE BOTH REGULATORY AND TRANSPORT ACTIVITIES.  
CC -1- SIMILARITY: Belongs to the ABC transporter family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
EMBL; L11870; AAA21933.1; -  
PIR; A36918; A36918.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR003439; ABC transporter.  
DR Pfam; PF00005; ABC tran; 1.  
DR ProDom; PD000006; ABC transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC TRANSPORTER\_1; 1.  
DR PROSITE; PS00893; ABC TRANSPORTER\_2; 1.  
KW ATP-binding.  
FT NP BIND 58 65 ATP (POTENTIAL).  
FT DOMAIN 238 259 ARG/LYS-RICH (BASIC).  
SQ SEQUENCE 308 AA; 34015 MW; 412023797D883E18 CRC64;  
  
Query Match 100.0%; Score 27; DB 1; Length 308;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDS 5  
|||||  
DB 258 GRGDS 262  
  
RESULT 9  
OSTP-RABIT STANDARD; PRT; 311 AA.  
ID OSTP-RABIT STANDARD; PRT; 311 AA.

Fri Apr 16 11:04:13 2004

us-09-991-588b-4.rsp

AC P31097; P46631;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein  
 GN 1) (SPP-1) (OC-1).  
 DE GN SPP1.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OC NCBI\_TaxID=9986;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Osteoclast;  
 RX MEDLINE=92360044; PubMed=1379809;  
 RA Tezuka K.-I., Sato T., Kamio H., Nijweide P.J., Tanaka K.,  
 RA Matsuo T., Ohta M., Kurihara N., Hakeda Y., Kamegawa M.;  
 RA "Identification of osteopontin in isolated rabbit osteoclasts.";  
 RL Biochem. Biophys. Res. Commun. 186:911-917(1992).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Myeloid;  
 RA Nasu K., Ishida T., Hijiya N., Setoguchi M., Akizuki S.,  
 RA Hijiya Y., Yamamoto S.;  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an  
 CC integral part of the mineralized matrix. Probably important to  
 CC cell-matrix interaction.  
 CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of  
 CC interferon-gamma and interleukin-12 and reducing production of  
 CC interleukin-10 and is essential in the pathway that leads to type  
 CC 2 immunity (By similarity).  
 CC -!- SUBUNIT: Ligand for integrin alpha-V/beta-3.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PTM: Extensively phosphorylated on serine residues.  
 CC -!- PTM: N- and O-glycosylated.  
 CC -!- SIMILARITY: Belongs to the osteopontin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D11411; BAA01993.1; -;  
 CC EMBL; J16544; BAA03980.1; -;  
 CC PIR; JC1191; JC1191.  
 CC InterPro; IPR002038; Osteopontin.  
 CC Pfam; PF00865; Osteopontin; 1.  
 CC PRINTS; PR00216; OSTEOPONTIN.  
 CC SMART; SM00017; OSTEO; 1.  
 CC PROSITE; PS00884; OSTEOPONTIN; 1.  
 CC Cytokine; Glycoprotein; Sialic acid; Biominalization; Cell adhesion;  
 CC Phosphorylation; Signal.  
 CC FT SIGNAL 1 16  
 CC BY SIMILARITY.  
 CC FT CHAIN 17 311  
 CC OSTEOPONTIN.  
 CC FT SITE 155 157  
 CC CELL ATTACHMENT SITE.  
 CC FT CARBOHYD 103 103  
 CC N-LINKED (GLCNAC...) (POTENTIAL).  
 CC FT CONFLICT 85 85  
 CC M -> V (IN REF. 2).  
 CC FT CONFLICT 100 100  
 CC R -> Q (IN REF. 2).  
 CC FT CONFLICT 128 128  
 CC E -> ESDE (IN REF. 2).  
 CC FT CONFLICT 151 151  
 CC T -> I (IN REF. 2).  
 CC SEQUENCE 311 AA; 35172 MW; 692750F63B7C85A5 CRC64;  
 SQ  
 Query Match 100.0%; Score 27; DB 1; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 |||||  
 Db 154 GRGDS 158

RESULT 10  
 OSTP\_HUMAN STANDARD; PRT; 314 AA.  
 ID AC P10451; Q15682; Q15683;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Osteopontin precursor (Bone sialoprotein 1) (Urinary stone protein)  
 DE (Secreted phosphoprotein 1) (SPP-1) (Nephropontin) (Uropontin).  
 GN SPP1 OR OPN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=89263749; PubMed=2726470;  
 RA Kiefer M.C., Bauer D.M., Barr P.J.;  
 RA "The cDNA and derived amino acid sequence for human osteopontin.";  
 RL Nucleic Acids Res. 17:3306-3306(1989).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=90353945; PubMed=1974876;  
 RA Young M.F., Kerr J.M., Termini J.D., Wewer U.M., Wang M.G.,  
 RA McBride O.W., Fisher L.W.;  
 RL "cDNA cloning, mRNA distribution and heterogeneity, chromosomal  
 RL location, and RFLP analysis of human osteopontin (OPN).";  
 RL Genomics 7:491-502(1990).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=92108068; PubMed=1729712;  
 RA Shiraga H., Min W., Vanbusen W.J., Clayman M.D., Miner D.,  
 RA Terrell C.H., Sherbotie J.R., Foreman J.W., Przysiecki C.,  
 RA Neilson E.G., Hoyer J.R.;  
 RL "Inhibition of calcium oxalate crystal growth in vitro by uropontin:  
 RL another member of the aspartic acid-rich protein superfamily.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:426-430(1992).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP Crosby A.H., Edwards S., Murray J.C., Dixon M.J.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RX MEDLINE=95031368; PubMed=7945249;  
 RA Hijiya N., Setoguchi M., Matsuura K., Hijiya Y., Akizuki S.,  
 RA Yamamoto S.;  
 RL "Cloning and characterization of the human osteopontin gene and its  
 RL promoter.";  
 RL Biochem. J. 303:255-262(1994).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RA Yu W., Sarginson J., Gibbs R.A.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.  
 RX MEDLINE=95139605; PubMed=7837791;  
 RA Saitoh Y., Kuratsu J., Takeshima H., Yamamoto S., Ushio Y.;  
 RL "Expression of osteopontin in human glioma: its correlation with the  
 RL malignancy.";  
 RL Lab. Invest. 72:55-63(1995).  
 RN [8]  
 RN SEQUENCE FROM N.A. (ISOFORM A).  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge C.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

DR GO:0003793; P:defense/immunity protein activity; TAS.  
 DR GO:0008083; P:growth factor activity; TAS.  
 DR GO:0005178; P:integrin binding; NAS.  
 DR GO:0006316; P:anti-apoptosis; ISS.  
 DR GO:0007267; P:cell-cell signaling; TAS.  
 DR GO:0007160; P:cell-matrix adhesion; NAS.  
 DR GO:0030595; P:immune cell chemotaxis; TAS.  
 DR GO:0006954; P:inflammatory response; TAS.  
 DR GO:0030502; P:negative regulation of bone mineralization; NAS.  
 DR GO:0001503; P:ossification; TAS.  
 DR GO:0004210; P:positive regulation of T-cell proliferation; TAS.  
 DR GO:0045637; P:regulation of myeloid blood cell differentiation; TAS.  
 DR GO:0042088; P:T-helper 1 type immune response; TAS.  
 DR InterPro: IPR002038; Osteopontin.  
 DR Pfam: PF00865; Osteopontin; 1.  
 DR PRINTS; PR00216; OSTEOPTNTN.  
 DR SMART; SM00017; OSTEO; 1.  
 DR PROSITE; PS00884; OSTEOPTNTN; 1.  
 KW Cytokine; Glycoprotein; Sialic acid; Biominalization; Cell adhesion;  
 KW Phosphorylation; Signal; Alternative splicing; Polymorphism.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 314 CELL ATTACHMENT SITE.  
 FT SITE 159 161 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 79 79 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 106 106 Missing (in isoform C).  
 FT VARSPLIC 31 57 /FTId=VSP\_003777.  
 FT VARSPLIC 58 71 Missing (in isoform B).  
 FT VARIANT 301 301 /FTId=VSP\_003778.  
 FT CONFLICT 188 188 R -> H (in dbSNP:4660).  
 FT CONFLICT 237 237 D -> H (IN REF. 7).  
 FT CONFLICT 275 278 T -> A (IN REF. 7).  
 FT CONFLICT 314 AA; 35422 MW; 49964292C4/52B86 CRC64; SHEF -> GMSL (IN REF. 2).  
 SQ SEQUENCE 100.0%; Score 27; DB 1; Length 314;  
 Query Match Best Local Similarity 100.0%; Pred. No. 56; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0;  
 Qy 1 GRGDS 5  
 Db 158 GRGDS 162  
 RESULT 11  
 OSTP RAT STANDARD; PRT; 317 AA.  
 ID AC P08721;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein 1) (SPP-1).  
 DE SPP1 OR SPP-1 OR 2B7.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]\_TaxID=10116;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=93054745; PubMed=1429723;  
 RA Singh K., Mukherjee A.B., de Vooge M.W., Mukherjee B.B.;  
 RT "Differential processing of osteopontin transcripts in rat kidney-  
 and osteoblast-derived cell lines";  
 RL J. Biol. Chem. 267:23847-23851(1992).  
 RN [2]  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=87067405; PubMed=3024151;  
 RA Oldberg A., Franzen A., Heinegaard D.;  
 RT "Cloning and sequence analysis of rat bone sialoprotein (osteopontin)  
 cDNA reveals an Arg-Gly-Asp cell-binding sequence.";

RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raka S.S., Lequellano N.A., Peters G.J., Abramson R.E., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fabey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Griwood J., Schmutz C., Myers R.M.,  
 RA Butterfield A.S., Klein J.E., Jones S.J.M., Marra M.A.,  
 RA Schnerch A., Schein J.R., Krzywinski X.L., Skalska U., Smallos D.E.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [9]  
 RP SEQUENCE OF 67-278 FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=92246377; PubMed=1575754;  
 RA Kohri K., Suzuki Y., Yoshida K., Yamamoto K., Anasaki N., Yamate T.,  
 RA Umekawa T., Iguchi M., Sinohara H., Kurita T.;  
 RT "Molecular cloning and sequencing of cDNA encoding urinary stone  
 protein, which is identical to osteopontin.";  
 RL Biochem. Biophys. Res. Commun. 184:859-864(1992).  
 CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an  
 integral part of the mineralized matrix. Probably important to  
 cell-matrix interaction.  
 CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of  
 interferon-gamma and interleukin-12 and reducing production of  
 interleukin-10 and is essential in the pathway that leads to type  
 I immunity (By similarity).  
 CC -!- SUBUNIT: Ligand for integrin alpha-V/beta-3.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=A; Synonyms=OPN-A, OP1B;  
 CC IsoId=P10451-1; Sequence=Displayed;  
 CC Name=B; Synonyms=OPN-B, OP1A;  
 CC IsoId=P10451-2; Sequence=VSP\_003778;  
 CC Name=C; Synonyms=OPN-C;  
 CC IsoId=P10451-3; Sequence=VSP\_003777;  
 CC -!- PTM: Extensively phosphorylated on serine residues.  
 CC -!- PTM: N- and C-glycosylated.  
 CC -!- DISEASE: This protein plays a principal role in urinary stone  
 formation as the stone matrix.  
 CC -!- SIMILARITY: Belongs to the osteopontin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; X13694; CAA31984.1; -  
 DR EMBL; J04765; AAA59974.1; -  
 DR EMBL; M83248; AAA17675.1; -  
 DR EMBL; U20758; AAA86886.1; -  
 DR EMBL; AF052124; AAC28619.1; -  
 DR EMBL; D28759; BAA05949.1; -  
 DR EMBL; D28760; BAA05950.1; -  
 DR EMBL; D28761; BAA05951.1; -  
 DR EMBL; D14813; BAA03554.1; -  
 DR EMBL; BC017387; AAH17387.1; -  
 DR FIC; S00028; S09575  
 DR GlycoSuiteDB; P10451.-  
 DR Genew; HGNC:11255; SPP1.  
 DR MIM; 166490; -  
 DR GO:0005578; C:extracellular matrix; TAS.  
 DR GO:0008189; F:apoptosis inhibitor activity; ISS.  
 DR GO:0042056; F:chemoattractant activity; TAS.  
 DR GO:0005125; F:cytokine activity; ISS.

Fri Apr 16 11:04:13 2004

us-09-991-588b-4.rsp

Proc. Natl. Acad. Sci. U.S.A. 83:8819-8823 (1986).

[3]  
SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; T=SSUE=Smooth muscle;  
RX MEDLINE=94013467; PubMed=8408622;  
RA Giachelli C.M., Bae N., Almeida M., Denhardt D.T., Alpers C.E.,  
RA Schwartz S.M.;  
RT "Osteopontin is elevated during neointima formation in rat arteries  
and is a novel component of human atherosclerotic plaques.";  
RL J. Clin. Invest. 92:1686-1696 (1993).  
CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an  
integral part of the mineralized matrix. Probably important to  
cell-matrix interaction.  
CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of  
interferon-gamma and interleukin-12 and reducing production of  
interleukin-10 and is essential in the pathway that leads to type  
I immunity (by similarity).  
CC -!- SUBUNIT: Ligand for integrin alpha-v/beta-3.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: Extensively phosphorylated on serine residues.  
CC -!- PTM: N- and O-glycosylated.  
CC -!- SIMILARITY: Belongs to the osteopontin family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M99252; AAA41765.1; -;  
EMBL; M14656; AAA41762.1; -;  
PIR; A25917; A25917.  
GO; GO:000575; C:extracellular matrix; ISS.  
GO; GO:0008189; F:apoptosis inhibitor activity; ISS.  
GO; GO:0042056; F:chemoattractant activity; ISS.  
GO; GO:0005125; F:cytokine activity; ISS.  
GO; GO:0003793; F:defense/immunity protein activity; ISS.  
GO; GO:0008093; F:growth factor activity; ISS.  
GO; GO:0005178; F:growth binding; ISS.  
GO; GO:0006916; P:anti-apoptosis; ISS.  
GO; GO:0007267; P:cell-cell signaling; ISS.  
GO; GO:0007150; P:cell-matrix adhesion; ISS.  
GO; GO:0030595; P:immune cell chemotaxis; ISS.  
GO; GO:0006954; P:inflammatory response; ISS.  
GO; GO:0030502; P:negative regulation of bone mineralization; ISS.  
GO; GO:0042102; P:positive regulation of T-cell proliferation; ISS.  
GO; GO:0045637; P:regulation of myeloid blood cell differenti. . .; ISS.  
GO; GO:0042088; P:T-helper 1 type immune response; ISS.  
InterPro: IPR002038; Osteopontin.  
Pfam: PF00865; Osteopontin; 1.  
PRINTS; PR00216; OSTEOPTNTN.  
SMART; SM00017; OSTEO; 1.  
PROSITE; PS00884; OSTEOPTNTN; 1.  
Cytokine; Glycoprotein; Sialic acid; Biomimetic; Cell adhesion;  
Phosphorylation; Signal.  
SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 327 OSTEOPTNTN.  
FT DOMAIN 86 96 POLY-ASP.  
FT SITE 144 146 CELL ATTACHMENT SITE.  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 8 8 F -> L (IN REF. 2).  
SQ SEQUENCE 317 AA; 34963 MW; 73CB5C21FF562310 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 317;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGFS 5  
DB 143 GRGDS 147

RESULT 12

FAR2\_HELAS STANDARD; FRT; 346 AA.

ID FAR2\_HELAS

AC P41870;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE FMRamide-related peptides type HF-1 precursor (TetraFARF).  
OS Helix aspersa (Brown garden snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Sigmurethra; Helicoidea; Helicidae; Helix.  
OX NCBI\_TaxID=6535;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ganglion;  
RC Lutz E.M., Macdonald M., Hettie S., Price D.A., Cottrell G.A.,  
RA Sommerville J.;  
RT "Structure of cDNA clones and genomic DNA encoding FMRamide-related  
peptides (FARPs) in Helix.";  
RL Mol. Cell. Neurosci. 3:373-382 (1992).  
RN [2]  
RP PARTIAL SEQUENCE (FLRF-AMIDE AND FMRF-AMIDE).  
RC TISSUE=Ganglion;  
RX MEDLINE=91116265; PubMed=1980513;  
RA Price D.A., Lesser M., Lee T.D., Dobie K.E., Greenberg M.J.;  
RT "Seven FMRamide-related and two SCP-related cardioactive peptides  
from Helix.";  
RL J. Exp. Biol. 154:421-437 (1990).  
CC -!- FUNCTION: Can function as both cardiorespiratory hormones and  
transmitters and may regulate cardiovascular function.  
CC -!- TISSUE SPECIFICITY: Central nervous system.  
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L20768; AAA57461.1; -;  
PIR; S77971; S77971.  
InterPro: IPR002544; FARP.  
Pfam; PF01581; FARP; 15.  
KW Neuropeptide; Amidation; Repeat; Signal;  
KW Cleavage on pair of basic residues.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 45  
FT PEPTIDE 47 51 QFYRF-AMIDE 1.  
FT PROPEP 54 94  
FT PEPTIDE 96 100 QFYRF-AMIDE 2.  
FT PROPEP 103 109  
FT PEPTIDE 112 115 FLRF-AMIDE 1.  
FT PROPEP 118 203  
FT PEPTIDE 206 209 FMRF-AMIDE 1.  
FT PEPTIDE 213 216 FMRF-AMIDE 2.  
FT PROPEP 219 226 FMRF-AMIDE 3.  
FT PEPTIDE 229 232 FMRF-AMIDE 4.  
FT PROPEP 235 243 FMRF-AMIDE 5.  
FT PEPTIDE 246 249 FMRF-AMIDE 6.  
FT PEPTIDE 253 256 FMRF-AMIDE 7.  
FT PROPEP 259 267 FMRF-AMIDE 8.  
FT PEPTIDE 270 273 FLRF-AMIDE 2.  
FT PROPEP 276 283 FMRF-AMIDE 6.  
FT PEPTIDE 287 290 FMRF-AMIDE 7.  
FT PROPEP 293 301 FMRF-AMIDE 8.  
FT PEPTIDE 304 307 FMRF-AMIDE 9.  
FT PROPEP 311 314  
FT PEPTIDE 317 325  
FT PROPEP 328 331

```

FT PEPTIDE 335 338 FMRF-AMIDE 10.
FT PROPEP 341 346
FT MOD_RES 51 51
FT MOD_RES 100 100
FT MOD_RES 115 115
FT MOD_RES 209 209
FT MOD_RES 216 216
FT MOD_RES 232 232
FT MOD_RES 249 249
FT MOD_RES 256 256
FT MOD_RES 273 273
FT MOD_RES 290 290
FT MOD_RES 307 307
FT MOD_RES 314 314
FT MOD_RES 331 331
FT MOD_RES 338 338
FT MOD_RES 346 AA; 41156 MW; E294F48B9A4FC734 CRC64;
SQ SEQUENCE 346 AA; 41156 MW; E294F48B9A4FC734 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 339 GRGDS 343

RESULT 13
ENP2_TORCA STANDARD; PRT; 350 AA.
ID ENP2_TORCA
AC 21401;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Electromotor neuron-associated protein 2 (Fragment).
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hymenoptera; Pteridaria; Batoida;
OC Torpediniformes; Torpedinidae; Torpedinidae; Torpedo.
OC NCBI_TaxID=1787;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=30091742; PubMed=2480872;
RX Ngsee J.K., Scheller R.H.;
RA "Isolation and characterization of two homologous cDNA clones from
RT torpedo electromotor neurons."
RL DNA 8:555-561(1989).
CC -!- SUBCELLULAR LOCATION: Associated with membranes of intracellular
CC organelles.
CC -!- SIMILARITY: TO ELECTROMOTOR NEURON-ASSOCIATED PROTEIN 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M30271; AAA49280.1; -.
CC F01; B33319; B33319.
CC KW Neurone.
CC FT NON_TER
CC SQ SEQUENCE 350 AA; 39201 MW; A1890D74AF136C6B CRC64;

Query Match 100.0%; Score 27; DB 1; Length 350;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 41 GRGDS 45

```

```

RESULT 14
COBL_MYCTU STANDARD; PRT; 390 AA.
ID COBL_MYCTU
AC Q10671;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Precorrin-6 C5,15-methyltransferase [decarboxylating] (EC 2.1.1.132)
DE (Precorrin-6 methyltransferase) (Precorrin-6Y methylase).
DE COBL OR RV2072C OR MT2132 OR WTCY49.11C.
GN Mycobacterium tuberculosis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295937; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., DeBoy R., Dodson R., Winn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson M.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- FUNCTION: CATALYZES THE METHYLATION OF BOTH C-5 AND C-15 IN
CC PRECORRIN-6Y TO FORM PRECORRIN-8X.
CC -!- CATALYTIC ACTIVITY: 2 S-adenosyl-L-methionine + precorrin-6Y = 2
CC S-adenosyl-L-homocysteine + precorrin-8X + CO(2).
CC -!- PATHWAY: Cobalamin biosynthesis.
CC -!- SIMILARITY: TO 8-TYPHIMURIUM CBIE; ALSO, LOW, TO OTHER
CC METHYLASES INVOLVED IN COBALAMIN BIOSYNTHESIS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Z73966; CAA98225.1; -.
CC F01; AEC07063; AAK46412.1; -.
CC TIGR; MT2132; -.
CC Tuberculin; RV2072C; -.
CC InterPro; IPR006365; Cobl.
CC InterPro; IPR000878; Cof/por Metransf.
CC InterPro; IPR000051; SAM_bind.
CC Pfam; PF00590; TP_methylase; 1.
CC TIGRfam; TIGR01468; cobl_cbl1; 1.
CC Cobalamin biosynthesis; Porphyrin biosynthesis; Transferase;
KW Methyltransferase; Complete proteome.
FT CONFLICT 327 327 D -> H (IN REF. 2).
SQ SEQUENCE 390 AA; 41854 MW; FB42EFF7562F21F3 CRC64;

```

Fri Apr 16 11:04:13 2004

Query Match 100.0%; Score 27; DB 1; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CQ 1 GRGS 5  
 DB 158 GRGS 162

RESULT 15

PRLA LYSEN STANDARD; PRT; 397 AA.  
 AC P00778;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Alpha-lytic protease precursor (EC 3.4.21.12; Alpha-lytic  
 endopeptidase).  
 GN ALPHA-LP.  
 OS Lysobacter enzymogenes.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Lysobacter.  
 OX NCBI\_TaxID=69;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 29487;  
 RC MEDLINE=89034140; PubMed=3053694;  
 RX Epstein D.M., Wensink P.C.;  
 RA "The alpha-lytic protease gene of Lysobacter enzymogenes. The  
 RT nucleotide sequence predicts a large prepro-peptide with homology to  
 RT pro-peptides of other chymotrypsin-like enzymes."  
 RL J. Biol. Chem. 263:16586-16590(1988).  
 RN [2]  
 RP REVISIONS.  
 RA Epstein D.M.;  
 RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 29487;  
 RC MEDLINE=89172068; PubMed=2334766;  
 RX Silen J.L., McGrath C.N., Smith K.R., Agard D.A.;  
 RA "Molecular analysis of the gene encoding alpha-lytic protease:  
 RT evidence for a preproenzyme."  
 RL Gene 69:237-244(1988).  
 RN [4]  
 RP SEQUENCE OF 199-396.  
 RC STRAIN=ATCC 29487;  
 RX MEDLINE=71039222; PubMed=5482494;  
 RA Olson M.O.J., Nagabushan N., Dzwiniel M., Smillie L.B.,  
 RA Whitaker D.R.;  
 RT "Purification and characterization of alpha-lytic protease: a bacterial homologue of  
 RT the pancreatic serine proteases."  
 RL Nature 228:438-442(1970).  
 RN [5]  
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RP MEDLINE=80074995; PubMed=117110;  
 RX Brayer G.D., Delbaere L.T.J., James M.N.G.;  
 RA "Molecular structure of the alpha-lytic protease from Myxobacter 495  
 RT at 2.8-A resolution."  
 RL J. Mol. Biol. 13:743-775(1979).  
 RN [6]  
 RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
 RP MEDLINE=86011557; PubMed=3300416;  
 RX Fujinaga M., Delbaere L.T.J., Brayer G.D., James M.N.G.;  
 RA "Refined structure of alpha-lytic protease at 1.7-A resolution.  
 RT Analysis of hydrogen bonding and solvent structure."  
 RL J. Mol. Biol. 184:479-502(1985).  
 RN [7]  
 RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 RP STRAIN=ATCC 29487;  
 RC MEDLINE=98393457; PubMed=9724517;  
 RX Peters R.J., Shiau A.K., Sohl J.L., Anderson D.E., Tang G.,

RA Silen J.L., Agard D.A.;  
 RT "Pro region C-terminus:protease active site interactions are critical  
 in catalyzing the folding of alpha-lytic protease."  
 RL Biochemistry 37:12058-12067(1998).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RX MEDLINE=99023190; PubMed=9808037;  
 RA Sauter N.K., Mao T., Rader S.D., Agard D.A.;  
 RT "Structure of alpha-lytic protease complexed with its pro region."  
 RL Nat. Struct. Biol. 5:945-950(1998).  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins, especially bonds  
 adjacent to L-alanine and L-valine residues in bacterial cell  
 walls, elastin and other proteins.  
 CC -!- SIMILARITY: Belongs to peptidase family S2A.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J04052; AAA25409.1; --  
 DR EMBL; M22763; AAA74111.1; --  
 DR PIR; A31772; TRYXB4.  
 DR PDB; 2ALP; 15-OCT-89.  
 DR PDB; 1GBA; 29-JAN-96.  
 DR PDB; 1GBB; 29-JAN-96.  
 DR PDB; 1GBC; 29-JAN-96.  
 DR PDB; 1GBD; 29-JAN-96.  
 DR PDB; 1GBE; 29-JAN-96.  
 DR PDB; 1GBF; 29-JAN-96.  
 DR PDB; 1GBH; 29-JAN-96.  
 DR PDB; 1GHI; 29-JAN-96.  
 DR PDB; 1GBJ; 29-JAN-96.  
 DR PDB; 1GBK; 29-JAN-96.  
 DR PDB; 1GBL; 29-JAN-96.  
 DR PDB; 1GBM; 29-JAN-96.  
 DR PDB; 2LPR; 15-JAN-93.  
 DR PDB; 3LPR; 15-JAN-93.  
 DR PDB; 5LPR; 15-JAN-93.  
 DR PDB; 6LPR; 15-JAN-93.  
 DR PDB; 7LPR; 15-JAN-93.  
 DR PDB; 8LPR; 15-JAN-93.  
 DR PDB; 9LPR; 15-JAN-93.  
 DR PDB; 1P01; 15-APR-90.  
 DR PDB; 1P02; 15-APR-90.  
 DR PDB; 1P03; 15-APR-90.  
 DR PDB; 1P04; 15-APR-90.  
 DR PDB; 1P05; 15-APR-90.  
 DR PDB; 1P06; 15-APR-90.  
 DR PDB; 1P09; 15-APR-90.  
 DR PDB; 1P10; 15-APR-90.  
 DR PDB; 1P11; 15-JAN-93.  
 DR PDB; 1P12; 15-JAN-93.  
 DR PDB; 1TAL; 01-APR-97.  
 DR PDB; 2JUL; 07-JUL-97.  
 DR PDB; 1BOQ; 12-AUG-98.  
 DR PDB; 1QQ4; 03-MAY-00.  
 DR PDB; 1QRM; 03-MAY-00.  
 DR PDB; 1ORX; 18-JUN-99.  
 DR PDB; 2PRO; 27-APR-99.  
 DR PDB; 3PRO; 27-APR-99.  
 DR PDB; 4PRO; 18-MAY-99.  
 DR MEROPS; S01.268; --  
 DR InterPro; IPR004236; AL\_protease.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001316; Peptidase S1e.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF02983; AL\_protease; 2.  
 DR Pfam; PF00089; trypsin; 1.

Fri Apr 16 11:04:13 2004

DR PRINTS; PRO0861; ALYTICPTASE.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; 3D-structure; Zymogen; Signal.  
 FT SIGNAL 1 24  
 FT PROPEP 25 199  
 FT CHAIN 200 397  
 FT DISULFID 216 236  
 FT DISULFID 300 310  
 FT DISULFID 336 369  
 FT ACT SITE 235 235  
 FT ACT SITE 262 262  
 FT ACT SITE 342 342  
 FT ACT SITE 358 358  
 FT CONFLICT 171 171  
 FT STRAND 201 204  
 FT TURN 205 205  
 FT STRAND 237 210  
 FT TURN 211 213  
 FT STRAND 214 217  
 FT STRAND 220 224  
 FT TURN 225 226  
 FT STRAND 227 232  
 FT HELIX 234 236  
 FT TURN 239 240  
 FT STRAND 242 245  
 FT TURN 246 247  
 FT STRAND 248 257  
 FT STRAND 260 260  
 FT STRAND 263 266  
 FT TURN 270 271  
 FT STRAND 273 280  
 FT TURN 281 282  
 FT STRAND 283 286  
 FT STRAND 289 289  
 FT TURN 295 296  
 FT STRAND 298 303  
 FT TURN 304 306  
 FT STRAND 307 322  
 FT TURN 323 324  
 FT STRAND 325 333  
 FT STRAND 338 338  
 FT TURN 339 340  
 FT TURN 342 343  
 FT STRAND 345 347  
 FT TURN 349 350  
 FT STRAND 352 360  
 FT TURN 364 365  
 FT STRAND 367 367  
 FT TURN 369 370  
 FT HELIX 373 375  
 FT STRAND 378 382  
 FT HELIX 384 390  
 FT TURN 391 391  
 FT STRAND 393 394  
 SQ SEQUENCE 397 AA; 41077 MW; 267FB6BFF57F33CB CRC64;

Query Match 100.0%; Score 27; DB 1; Length 397;  
 Best local Similarity 100.0%; Pred. No. 73;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDS 5  
 Db 338 GRGDS 342

Search completed: April 16, 2004, 07:22:11  
 Job time : 9.75 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 39.5 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588b-4

Perfect score: 27

Sequence: 1 GRGDS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2300000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvivirus.\*
- 16: sp\_bacterioph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	23	4 Q9UC00	Q9UC00 homo sapien
2	27	100.0	60	16 Q83CJ4	Q83CJ4 coxiella bu
3	27	100.0	65	6 Q95M28	Q95M28 ovis aries
4	27	100.0	65	6 Q95M46	Q95M46 bos taurus
5	27	100.0	83	16 Q7MGA3	Q7MGA3 bordetella
6	27	100.0	83	16 Q7MGA7	Q7MGA7 bordetella
7	27	100.0	83	16 Q7VU91	Q7VU91 bordetella
8	27	100.0	84	10 Q8H522	Q8H522 oryza sativ
9	27	100.0	86	2 Q9E2D2	Q9E2D2 rhizobium m
10	27	100.0	105	6 Q9TSY3	Q9TSY3 sus scrofa
11	27	100.0	116	16 Q914B0	Q914B0 pseudomonas
12	27	100.0	125	5 Q8SW51	Q8SW51 encephalito
13	27	100.0	125	10 Q9AWZ1	Q9AWZ1 oryza sativ
14	27	100.0	134	16 Q9RYM3	Q9RYM3 deinococcus
15	27	100.0	141	11 Q9CVR5	Q9CVR5 mus musculu
16	27	100.0	143	16 Q88XW7	Q88XW7 lactobacill

Q8Z5E5 salmonella  
Q8X1S5 emericeella  
Q8X1S7 emericeella  
Q8Y6Y7 emericeella  
Q9VB29 drosophila  
Q9FHX7 arabidopsis  
Q8RWR9 arabidopsis  
Q8VBT7 rattus norv  
Q876Z0 ashbya goss  
Q8HAG3 homo sapien  
Q8Z5C5 streptomyce  
Q7WXB1 alcaligenes  
Q8E8A4 rhizobium  
Q8VZT0 arabidopsis  
Q81G22 arabidopsis  
Q84WW0 arabidopsis  
Q82K28 rhizobium m  
Q91VH4 mus musculu  
Q22136 arabidopsis  
Q9BH02 homo sapien  
Q8E715 arabidopsis  
Q8C987 mus musculu  
Q8F21 streptomyce  
Q81Z6 rhizobium i  
Q8E8G4 shewanella  
Q91832 coturnix co  
Q64125 mus sp. ear  
Q8FTD5 corynebacte  
Q8NBK2 homo sapien

# ALIGNMENTS

## RESULT 1

Q9UC00 PRELIMINARY; PRT; 23 AA.  
AC Q9UC00;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Enhancement of wound HEALING process.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95130623; PubMed=7829572;  
RA Pierschbacher M.D., Polarek J.W., Craig W.S., Tschopp J.F.,  
RA Sipes N.J., Harper J.R.;  
RL J. Cell. Biochem. 56:150-154(1994).  
DR GO:0009611; P:response to wounding; TAS.  
SQ SEQUENCE 23 AA; 228 MW; CE73999CB9903891 CRC64;

Query Match 100.0%; Score 27; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
Db 1 GRGDS 5

## RESULT 2

Q83CJ4 PRELIMINARY; PRT; 60 AA.  
AC Q83CJ4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN CBU1121.

Fri Apr 16 11:04:14 2004

```

OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coccidiaceae; Coxiella.
OX NCB1_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Name Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=2704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward K.L., Tettelin H., Daviden T.M., Beaman M.J.,
RA DeBoy R.T., Dougherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
RT burnetii.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
RL EMBL; AB016963; AAC90634.1; -.
DR TIGR; CBU1121; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 60 AA; 6974 MW; 82D9D09533B9C9F2 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 GRGDS 5
DB 56 GRGDS 60

RESULT 3
Q95M28 PRELIMINARY; PRT; 65 AA.
ID Q95M28
AC Q95M28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Tight junction protein 2 (Fragment).
GN ZO2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCB1_TaxID=3940;
RN [1]
RP SEQUENCE FROM N.A.
RC Pfaffl M.W., Dzidic A., Rojas P.G., Bruckmaier R.M., Schams D.;
RT "Effects of an induced mammosgenesis and lactogenesis in sheep on the
RT mRNA expression levels of immune globulin receptors (FcRn; pIGR) and
RT zona occludens proteins (Occludin, ZO1; ZO2; ZO3).";
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ313187; CAC4191.1; -.
DR EMBL; AJ313187; CAC4191.1; -.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7460 MW; 7EB5829EBE641EPB4 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 GRGDS 5
DB 50 GRGDS 54

RESULT 4
Q95M46 PRELIMINARY; PRT; 65 AA.
ID Q95M46
AC Q95M46;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Tight junction protein 2 (Fragment).
GN ZO2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCB1_TaxID=3940;
RN [1]
RP SEQUENCE FROM N.A.
RC Pfaffl M.W., Dzidic A., Rojas P.G., Bruckmaier R.M., Schams D.;
RT "Effects of an induced mammosgenesis and lactogenesis in sheep on the
RT mRNA expression levels of immune globulin receptors (FcRn; pIGR) and
RT zona occludens proteins (Occludin, ZO1; ZO2; ZO3).";
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ313187; CAC4191.1; -.
DR EMBL; AJ313187; CAC4191.1; -.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7460 MW; 7EB5829EBE641EPB4 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 GRGDS 5
DB 50 GRGDS 54

RESULT 5
Q7WGA3 PRELIMINARY; PRT; 83 AA.
ID Q7WGA3
AC Q7WGA3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative cell division topological specificity factor.
GN MINE OR BB4016.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCB1_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STEAIN=RB50 / ATCC BAA-588;
RC MEDLINE=22827954; PubMed=12910271;
RX Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RX Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RX Cerdano-Farraga A.M., Temple L., James K., Harrison N., Cherevach I.,
RX Achtman M., Atkin R., Baker S., Basham D., Bason N., Doggett J.,
RX Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RX Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RX Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RX Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RX Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RX Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640449; CAE34379.1; -.
KW Cell division; Complete proteome.
SQ SEQUENCE 83 AA; 9489 MW; 98431A7ABDD2F58 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 GRGDS 5
DB 31 GRGDS 35

RESULT 6
Q7W4S7 PRELIMINARY; PRT; 83 AA.
ID Q7W4S7

```

Fri Apr 16 11:04:14 2004

```

AC Q794S7;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative cell division topological specificity factor.
DE MINE OR BP3229.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson M.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RW EMBL; BX640434; CA38865.1; -
KW Cell division; Complete proteome.
SQ SEQUENCE 83 AA; 9489 MW; 98431A7A7BDD2F58 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 31 GRGDS 35

RESULT 7
ID Q7VU91 PRELIMINARY; PRT; 83 AA.
AC Q7VU91;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative cell division topological specificity factor.
DE MINE OR BP3229.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson M.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RW EMBL; BX640420; CA53495.1; -
KW Cell division; Complete proteome.
SQ SEQUENCE 83 AA; 9489 MW; 98431A7A7BDD2F58 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 31 GRGDS 35

RESULT 7
ID Q7VU91 PRELIMINARY; PRT; 83 AA.
AC Q7VU91;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative cell division topological specificity factor.
DE MINE OR BP3229.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson M.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RW EMBL; BX640420; CA53495.1; -
KW Cell division; Complete proteome.
SQ SEQUENCE 83 AA; 9489 MW; 98431A7A7BDD2F58 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 31 GRGDS 35

RESULT 8
ID Q8H5Z2 PRELIMINARY; PRT; 84 AA.
AC Q8H5Z2;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE QJ1003 H02.1 protein.
DE QJ1003 H02.1.
GN QJ1003_H02.1.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:QJ1003 H02.1";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP003700; BAC21306.1; -
SQ SEQUENCE 84 AA; 8967 MW; EE6721C599FE6401 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 16 GRGDS 20

RESULT 9
ID Q9EZD2 PRELIMINARY; PRT; 86 AA.
AC Q9EZD2;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
DE Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RU11/001;
RA Muschler P.F.J., Endter C.;
RT "Isolation of mcp genes from Sinorhizobium meliloti.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF298190; AAG37856.1; -
GN Hypothetical protein.
SQ SEQUENCE 86 AA; 9128 MW; 87C5BF56B176DD2A CRC64;

Query Match 100.0%; Score 27; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5
DB 35 GRGDS 39

```

Fri Apr 16 11:04:14 2004

us-09-991-588b-4.rspt

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10  
Q9TSY3  
ID Q9TSY3 PRELIMINARY; PRT; 105 AA.  
AC Q9TSY3  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Osteopontin; (Fragment).  
GN SPPL.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=323;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=10612243;  
RX Knoll A., Stratil A., Cepica S., Dvorak J.;  
RA "Length polymorphism in an intron of the porcine osteopontin (SPPL)  
RT gene is caused by the presence or absence of a SINE (PFR-1) element."  
RL Anim. Genet. 30:466-466(1999).  
DR EMBL; AJ237667; CAB61259.1; -.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0001503; P:ossification; IEA.  
DR InterPro; IPR002038; Osteopontin.  
DR Pfam; PF009545; Osteopontin; 1.  
DR PRINTS; PR02216; OSTEOPONTIN.  
FT NON TER 1  
FT NON TER 105  
SQ SEQUENCE 105 AA; 11489 MW; F2EPCBAE560773CB CRC64;

Query Match 100.0%; Score 27; DB 6; Length 105;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
Db 24 GRGDS 28

RESULT 11  
Q9I4B0  
ID Q9I4B0 PRELIMINARY; PRT; 116 AA.  
AC Q9I4B0  
DT 01-MAY-2001 (TrEMBLrel. 16, Created)  
DT 01-MAY-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein: PA1228.  
GN PA1228.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Garber R.L., Goltzer L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
Brody L.L., Coulter S.N., Folger K.R., Kas A., Labig K., Lim R.M.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL; AE004552; AAG04617.1; -.  
DR PIR; C83492; C83492.  
DR Hypothetical protein; Complete proteome.  
SQ SEQUENCE 116 AA; 12362 MW; 9FBA2C09CFF0076A CRC64;

Query Match 100.0%; Score 27; DB 16; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
Db 15 GRGDS 19

RESULT 12  
Q8SW51  
ID Q8SW51 PRELIMINARY; PRT; 125 AA.  
AC Q8SW51  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similarity with homeobox domain-containing protein.  
GN ECUC3 0600.  
OS Encephalitozoon cuniculi.  
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
OX NCBI\_TaxID=6035;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GB-M1;  
RA Genoscope;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GB-M1;  
RX MEDLINE=21576510; PubMed=11719806;  
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
Prenier G., Barbe V., Peyretallade E., Brottier P., Wincker P.,  
Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
Weissenbach J., Vivares C.P.;  
RT "Genome sequence and gene compaction of the eukaryote parasite  
Encephalitozoon cuniculi."  
RL Nature 414:450-453(2001).  
DR EMBL; AL590443; CAD26206.1; -.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS50071; HOMEBOX 2; 1.  
DR Homeobox; DNA-binding; Nuclear protein.  
SQ SEQUENCE 125 AA; 14342 MW; 6CCD4F3EA61C1309 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 125;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDS 5  
Db 24 GRGDS 28

RESULT 13  
Q9AWZ1  
ID Q9AWZ1 PRELIMINARY; PRT; 125 AA.  
AC Q9AWZ1  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE P0492F05.21 protein (P0443E07.13 protein).  
GN P0492F05.21 OR P0443E07.13.  
OS Oryza sativa (Rice); and  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530, 39947;  
[1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=O.sativa; STRAIN=cv. Nipponbare;

Fri Apr 16 11:04:14 2004

RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0492F05."; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0492F05."; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF002902; BAB32719.1; -.  
 DR EMBL; AF002900; BAB92110.1; -.  
 DR Genbank; Q9KXZ1; -.  
 SQ SEQUENCE 125 AA; 13576 MW; 461B09A9F45ECF56 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 27 GRGDS 31

## RESULT 14

Q9RYM3 PRELIMINARY; PRT; 134 AA.  
 AC Q9RYM3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein DRA0289.  
 GN DRA0289.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R1 / ATCC 13939 / DSM 23539 / NCIB 9279;  
 RX MEDLINE=2036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.V., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.B., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans R1.";  
 RJ Science 286:1571-1577(1999).  
 DR EMBL; AE001863; AAF12529.1; -.  
 DR PIR; F75582; F75582.  
 DR TIGR; DRA0289; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 134 AA; 13975 MW; 6DEBF776F01F9C5B CRC64;

Query Match 100.0%; Score 27; DB 16; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 128 GRGDS 132

## RESULT 15

Q9CVR5 PRELIMINARY; PRT; 141 AA.  
 AC Q9CVR5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 1700062622Rik protein (Fragment).  
 GN 1700062622Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,  
 RA Sakai K., Oikido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK006863; BAB24770.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 141 AA; 14663 MW; 594AEC29E4A6277 CRC64;

Query Match 100.0%; Score 27; DB 11; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDS 5  
 DB 63 GRGDS 67

Search completed: April 16, 2004, 07:27:44  
 Job time : 41.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:28:08 ; Search time 26.9 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588B-6

Perfect score: 21

Sequence: 1 KGDS 4

Scoring table: BLOSUM62

Gapop -0.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	11	US-09-991-588B-6
2	21	100.0	10	9	US-09-308-511-2
3	21	100.0	10	12	US-10-351-630-2
4	21	100.0	14	9	US-09-885-441-6
5	21	100.0	14	12	US-10-424-836-6
6	21	100.0	18	14	US-10-225-567A-1779
7	21	100.0	19	14	US-10-083-357-1141
8	21	100.0	21	14	US-10-258-105-51
9	21	100.0	24	9	US-09-885-441-9
10	21	100.0	24	12	US-10-424-836-9
11	21	100.0	24	14	US-10-258-105-43
12	21	100.0	26	9	US-09-864-761-36454
13	21	100.0	26	9	US-09-864-761-45482
14	21	100.0	32	12	US-10-424-599-206472
15	21	100.0	35	9	US-09-885-441-3

16	21	100.0	35	12	US-10-424-836-3	Sequence 3, Appli
17	21	100.0	35	14	US-10-029-386-34275	Sequence 34275, A
18	21	100.0	37	14	US-10-029-386-31407	Sequence 31407, A
19	21	100.0	39	9	US-09-864-761-41933	Sequence 41933, A
20	21	100.0	42	12	US-10-424-599-256020	Sequence 256020,
21	21	100.0	42	14	US-10-321-857-133	Sequence 133, App
22	21	100.0	42	14	US-10-318-675-133	Sequence 133, App
23	21	100.0	43	12	US-10-424-599-274624	Sequence 274624,
24	21	100.0	43	14	US-10-321-857-96	Sequence 96, Appl
25	21	100.0	43	14	US-10-318-675-96	Sequence 96, Appl
26	21	100.0	44	9	US-09-975-143-17	Sequence 17, Appl
27	21	100.0	44	12	US-10-424-599-181462	Sequence 181462,
28	21	100.0	45	14	US-10-258-105-11	Sequence 11, Appl
29	21	100.0	45	14	US-10-258-105-52	Sequence 52, Appl
30	21	100.0	45	14	US-10-321-857-64	Sequence 64, App
31	21	100.0	45	14	US-10-321-857-120	Sequence 120, App
32	21	100.0	45	14	US-10-318-675-120	Sequence 120, App
33	21	100.0	46	12	US-10-424-599-169232	Sequence 169232,
34	21	100.0	46	12	US-10-424-599-243448	Sequence 243448,
35	21	100.0	46	12	US-10-424-599-243448	Sequence 243448,
36	21	100.0	47	9	US-09-864-761-44927	Sequence 44927, A
37	21	100.0	47	12	US-10-424-599-284023	Sequence 284023,
38	21	100.0	48	14	US-10-258-105-42	Sequence 42, Appl
39	21	100.0	48	14	US-10-258-105-44	Sequence 44, Appl
40	21	100.0	56	12	US-10-424-599-149707	Sequence 149707,
41	21	100.0	57	9	US-09-925-299-1020	Sequence 1020, Ap
42	21	100.0	57	10	US-09-925-299-1020	Sequence 1020, Ap
43	21	100.0	58	12	US-10-424-599-276892	Sequence 276892,
44	21	100.0	59	12	US-10-424-599-239819	Sequence 239819,
45	21	100.0	60	10	US-09-899-495-80	Sequence 80, Appl

## ALIGNMENTS

RESULT 1  
US-09-991-588B-6  
; Sequence 6, Application US/09991588B  
; Publication No. US20030219429A1  
; GENERAL INFORMATION:  
; APPLICANT: Budy, John A.  
; TITLE OF INVENTION: Compositionand Method for Bone Regeneratio.c  
; FILE REFERENCE: 1008-120, US  
; CURRENT APPLICATION NUMBER: US/09/991.588B  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 09/122,348  
; PRIOR FILING DATE: 1998-07-24  
; NUMBER OF SEQ ID NOS: 23  
; SEQ ID NO 6  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Purchased commercially or sequence is synthesized  
US-09-991-588B-6

Query Match 100.0%; Score 21; DB 11; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KGDS 4  
Db 1 KGDS 4

RESULT 2  
US-09-308-511-2  
; Sequence 2, Application US/09308511  
; Patent No. US20020103145A1  
; GENERAL INFORMATION:  
; APPLICANT: BOT and BONA  
; TITLE OF INVENTION: IMMUNIZATION OF INFANTS  
; NUMBER OF SEQUENCES: 20

us-09-991-588b-6.rapb

Fri Apr 16 11:04:14 2004

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue &
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,511
FILING DATE: 19-May-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: 29889-165/29528
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Influenza Virus
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: HAI hemagglutinin protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-308-511-2
Query Match 100.0%; Score 21; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 5 KGDS 8

RESULT 3
US-10-351-630-2
Sequence 2, Application US/10351630
Patent No. US20040053871A1
GENERAL INFORMATION:
APPLICANT: BOT and BONA
TITLE OF INVENTION: IMMUNIZATION OF INFANTS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue &
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

```

```

SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/351,630
FILING DATE: 24-Jan-2003
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: 29889-165/29528
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Influenza Virus
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: HAI hemagglutinin protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-351-630-2
Query Match 100.0%; Score 21; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 5 KGDS 8

RESULT 4
US-09-885-441-6
Sequence 6, Application US/09885441
Patent No. US20020146407A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Eosinophil Serine
PROTEASE-1-LIKE ENZYME
FILE REFERENCE: 04974.00512
CURRENT APPLICATION NUMBER: US/09/885,441
CURRENT FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/212,844
PRIOR FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: US 60/244,171
PRIOR FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: US 60/279,766
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: PCT/
PRIOR FILING DATE: 2001-06-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: BLOCKS domain
US-09-885-441-6
Query Match 100.0%; Score 21; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 KGDS 4
      ||||
Db      1 KGDS 4

RESULT 7
US-10-083-357-1141
; Sequence 1141, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiangdong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 1141
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1141

Query Match      100.0%; Score 21; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KGDS 4
      ||||
Db      5 KGDS 8

RESULT 8
US-10-258-105-51
; Sequence 51, Application US/10258105
; Publication No. US20030158382A1
; GENERAL INFORMATION:
; APPLICANT: Wakamiya et al.
; TITLE OF INVENTION: No. US20030158382A1el Collectin
; FILE REFERENCE: 19036/38785
; CURRENT APPLICATION NUMBER: US/10/258,105
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/JP01/03468
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: JP 2000-120358
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 51
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-258-105-51

Query Match      100.0%; Score 21; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KGDS 4
      ||||
Db      3 KGDS 6

RESULT 9
US-09-885-441-9
; Sequence 9, Application US/09885441
; Patent No. US20020146407A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; FILE REFERENCE: 04974.00512
; FILE REFERENCE: Protease-1-Like Enzyme
; CURRENT APPLICATION NUMBER: US/09/885,441
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844

US-10-424-836-6
; Sequence 6, Application US/10424836
; Publication No. US2003024430A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; FILE REFERENCE: 04974.00512
; FILE REFERENCE: Protease-1-Like Enzyme
; CURRENT APPLICATION NUMBER: US/10/424,836
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US/09/885,441
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/244,171
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: US 60/279,766
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BLOCKS domain
US-10-424-836-6

Query Match      100.0%; Score 21; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KGDS 4
      ||||
Db      5 KGDS 8

RESULT 6
US-10-225-567A-1779
; Sequence 1779, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Reusch, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1779
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-1779

Query Match      100.0%; Score 21; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



us-09-991-588b-6.rapb

Fri Apr 16 11:04:14 2004

```

; FILE REFERENCE: 19036/38785
; CURRENT APPLICATION NUMBER: US/10/258,105
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/JP01/03468
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: JP 2000-120358
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 43
; TYPE: PRT
; LENGTH: 24
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence of a collagen-like domain of mutated
; OTHER INFORMATION: No. US20030158382A1el Collectin.
; US-10-258-105-43

Query Match 100.0%; Score 21; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 6 KGDS 9

RESULT 12
US-09-864-761-36454
; Sequence 36454, Application US/03864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Ranzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.5
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/03661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408

; FILE REFERENCE: 19036/38785
; CURRENT APPLICATION NUMBER: US/10/258,105
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/JP01/03468
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: JP 2000-120358
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 43
; TYPE: PRT
; LENGTH: 24
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence of a collagen-like domain of mutated
; OTHER INFORMATION: No. US20030158382A1el Collectin.
; US-10-258-105-43

Query Match 100.0%; Score 21; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 4 KGDS 7

RESULT 10
US-10-424-836-9
; Sequence 9, Application US/10424836
; Publication No. US2003022430A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; TITLE OF INVENTION: Protease-1-Like Enzyme
; FILE REFERENCE: 04974.00512
; CURRENT APPLICATION NUMBER: US/10/424,836
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US/09/885,441
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/244,171
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: US 60/279,766
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BLOCKS domain
; US-10-424-836-9

Query Match 100.0%; Score 21; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 4 KGDS 7

RESULT 11
US-10-258-105-43
; Sequence 43, Application US/10258105
; Publication No. US20030158382A1
; GENERAL INFORMATION:
; APPLICANT: Wakamiya et al.
; TITLE OF INVENTION: No. US20030158382A1el Collectin

```

Query Match	Score	DB	Length	Indels	Gaps
Best Local Similarity	100.0%	21	26	0	0
Matches	4	Conservative	0	Mismatches	0
QY	1 KGDS	4			
DB	9 KGDS	12			
<p>US-09-864-761-45482</p> <p>Query Match 100.0%; Score 21; DB 9; Length 26;</p> <p>Best Local Similarity 100.0%; Pred. No. 4e+02;</p> <p>Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>					
<p>RESULT 13</p> <p>US-09-864-761-45482</p> <p>Sequence 45482, Application US/09864761</p> <p>Patent No. US20020048763A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Penn, Sharon G.</p> <p>APPLICANT: Rank, David R.</p> <p>APPLICANT: Hanzel, David K.</p> <p>APPLICANT: Chen, Wensheng</p> <p>TITLE OF INVENTION: HUMAN GENE EXPRESSION ANALYSIS BY MICROARRAY</p> <p>FILE REFERENCE: Aescmca-X-1</p> <p>CURRENT APPLICATION NUMBER: US/09/864,761</p> <p>CURRENT FILING DATE: 2001-05-23</p> <p>PRIOR APPLICATION NUMBER: US 60/180,312</p> <p>PRIOR FILING DATE: 2000-02-04</p> <p>PRIOR APPLICATION NUMBER: US 60/207,456</p> <p>PRIOR FILING DATE: 2000-05-26</p> <p>PRIOR APPLICATION NUMBER: US 09/532,366</p> <p>PRIOR FILING DATE: 2000-08-03</p> <p>PRIOR APPLICATION NUMBER: GB 24263.6</p> <p>PRIOR FILING DATE: 2000-10-04</p> <p>PRIOR APPLICATION NUMBER: US 60/236,359</p> <p>PRIOR FILING DATE: 2000-09-27</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00666</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00667</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00664</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00669</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00665</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00668</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00663</p> <p>PRIOR FILING DATE: 2001-01-30</p> <p>PRIOR APPLICATION NUMBER: PCT/US01/00662</p>					

us-09-991-588b-6.rapb

Fri Apr 16 11:04:14 2004

```

; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; FILE OF INVENTION: Protease-1-Like Enzyme
; CURRENT APPLICATION NUMBER: US 09/885,441
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/244,171
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: US 60/279,766
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BLOCKS domain
US-09-885-441-3

```

```

Query Match      100.0%; Score 21; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 KGDS 4
      ||||
Db     12 KGDS 15

```

Search completed: April 16, 2004, 08:03:16  
Job time : 26.9 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 8.4 seconds  
(without alignments)

45.805 Million cell updates/sec

Title: US-09-991-588B-6

Perfect score: 2

Sequence: 1 KGDS 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	21	100.0	43	2	T07124	RNA-directed DNA p
2	21	100.0	53	2	AF1734	hypothetical prote
3	21	100.0	59	2	B90990	hypothetical prote
4	21	100.0	63	2	AG3289	hypothetical prote
5	21	100.0	65	2	H72805	sp51 protein - Myc
6	21	100.0	75	2	E90974	probable cell divi
7	21	100.0	76	2	T09476	hypothetical prote
8	21	100.0	77	2	D81267	5S ribosomal prot
9	21	100.0	78	2	H90114	putative small nuc
10	21	100.0	79	2	I55584	complement C4A - h
11	21	100.0	89	2	E69972	conserved hypothet
12	21	100.0	94	2	D82712	histone-like prote
13	21	100.0	94	2	T00965	hypothetical prote
14	21	100.0	96	1	LECE29	hypothetical prote
15	21	100.0	98	2	I40616	DNA-binding protei
16	21	100.0	99	2	AH1886	hypothetical prote
17	21	100.0	101	2	A28487	apolipoprotein C-I
18	21	100.0	103	2	AB3348	LSU ribosomal prot
19	21	100.0	103	2	S06386	serum amyloid A pr
20	21	100.0	104	2	S15395	tissue kallikrein-
21	21	100.0	104	2	E97300	hypothetical prote
22	21	100.0	104	2	AC2029	probable lipoprote
23	21	100.0	105	2	AC2943	pectin degradation
24	21	100.0	105	2	G98339	pectin degradation
25	21	100.0	111	2	G91043	detox protein [imp
26	21	100.0	111	2	B85888	detox protein [imp
27	21	100.0	111	2	H65020	cchA protein - Esc
28	21	100.0	111	2	F96590	hypothetical prote
29	21	100.0	111	2	AC1648	hypothetical prote

30	21	100.0	113	2	E90345	hypothetical prote
31	21	100.0	114	2	E69137	hypothetical prote
32	21	100.0	118	2	H71640	hypothetical prote
33	21	100.0	120	2	F69934	hypothetical prote
34	21	100.0	124	2	S13199	profilin p - slime
35	21	100.0	125	2	B35273	profilin p - slime
36	21	100.0	125	2	AB1747	glycine cleavage s
37	21	100.0	125	2	AI1377	glycine cleavage s
38	21	100.0	125	2	E87514	hypothetical prote
39	21	100.0	126	2	A97110	hypothetical prote
40	21	100.0	128	2	C69183	hypothetical prote
41	21	100.0	130	2	S66474	cytochrome P450 (C
42	21	100.0	131	2	S66475	cytochrome P450 (C
43	21	100.0	133	1	VKLJCE	trans-regulatory s
44	21	100.0	134	2	D90956	probable transcript
45	21	100.0	134	2	A85805	unknown protein en

## ALIGNMENTS

### RESULT 1

T07124  
RNA-directed DNA polymerase (EC 2.7.7.49) (clone DES6) - potato Ty-copia-like retrotrans  
N:Alternate names: reverse transcriptase  
C:Species: Solanum tuberosum (potato)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T07124  
R:Kuipers, A.G.J.  
submitted to the EMBL Data Library, April 1998  
A:Reference number: Z15935  
A:Accession: T07124  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-43 <XUI>  
A:Cross-references: EXBL:AJ228807; NID:e1288232; P:EN:CAA13064.1; PID:e1288233  
A:Experimental source: cv. Desiree; dev. stage full grown; young leaves  
C:Genetics:  
A:Mobile element: Ty-copia-like retrotransposon  
C:Function:  
A:Description: catalyzes the extension of a DNA strand at the 3'end by one nucleotide at  
C:Keywords: nucleotidyltransferase

Query Match 100.0%; Score 21; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||  
DB 29 KGDS 32

### RESULT 2

AF1734  
hypothetical protein lin2419 [imported] - Listeria innocua (strain Clp11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1734  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H  
.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1734  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-53 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CACS97646.1; PID:gi6414941; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:

C:Species: Mycobacterium phage D29  
C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 20-Apr-2001  
C:Accession: H72805  
R:Ford, M.B.; Sarkis, G.J.; Belanger, A.E.; Hendrix, R.W.; Hatfull, G.F.  
J. Mol. Biol. 279, 143-164, 1998  
A:Title: Genome structure of mycobacteriophage D29: Implications for phage evolution.  
A:Reference number: A72800; PMID:98300335; PMID:9636706  
A:Accession: H72805  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-65 <FOR>  
A:Cross-references: GB:AF022214; NID:G5172250; PIDN:AA018491.1; PID:G5172298  
C:Genetics:  
A:Gene: 51

Query Match 100.0%; Score 21; DB 2; Length 65;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||||  
Db 48 KGDS 51

RESULT 6  
E90974  
probable cell division control protein [imported] - Escherichia coli (strain O157:H7, s  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: E90974  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; PMID:11258796  
A:Accession: E90974  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-75 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA036188.1; PID:G13362233; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECS2765

Query Match 100.0%; Score 21; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||||  
Db 72 KGDS 75

RESULT 7  
T09476  
hypothetical protein - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C:Accession: T09476  
R:Murphy, L.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z16683  
A:Accession: T09476  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-76 <MUR>  
A:Cross-references: EMBL:AL021546; NID:el248287; PID:el248289  
C:Genetics:  
A:Gene: 15E1.1  
A:Map position: 12  
A:Introns: 49/3

Query Match 100.0%; Score 21; DB 2; Length 76;

RESULT 5  
H72805  
gp51 protein - Mycobacterium phage D29

A:Gene: lin2413

Query Match 100.0%; Score 21; DB 2; Length 53;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||||  
Db 35 KGDS 38

RESULT 3  
E90990  
hypothetical protein ECS2890 [imported] - Escherichia coli (strain O157:H7, substrain R  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: B90990  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99623; PMID:11258796  
A:Accession: B90990  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-59 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA036188.1; PID:G13362359; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECS2890

Query Match 100.0%; Score 21; DB 2; Length 59;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||||  
Db 38 KGDS 41

RESULT 4  
AG3289  
hypothetical protein BMEI0301 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AG3289  
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AG3289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-63 <MUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51482.1; PID:G17982194; GSPDB:GN03190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0301  
A:Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
|||||  
Db 51 KGDS 54

Fri Apr 16 11:04:15 2004

us-09-991-588b-6.rpr

C;Accession: I55584  
 R;Barba, G.; Rittner, C.; Schneider, P.M.  
 J. Clin. Invest. 91, 1681-1686, 1993  
 A;Title: Genetic basis of human complement C4A deficiency. Detection of a point mutation  
 A;Reference number: I55584; MUID:93232295; PMID:6473511  
 A;Accession: I55584  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-79 <RES>  
 A;Cross-references: GB:58715; NID:G385258; PIDN:AB26211.1; PID:G385259  
 C;Genetics:  
 A;Gene: GDB:C4A  
 A;Cross-references: GDB:119732; OMIM:120810  
 A;Map position: 6p21.3-6p21.3  
 A;Introns: 77/2

Query Match 100.0%; Score 21; DB 2; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 29 KGDS 32

RESULT 8  
 D81267  
 50S ribosomal protein L24 Cj1696c [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C;Species: Campylobacter jejuni  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C;Accession: D81267  
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajadaram, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A;Reference number: A81250; MUID:20150912; PMID:10688204  
 A;Accession: D81267  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-77 <PAR>  
 A;Cross-references: GB:A0139079; GB:AL111168; NID:G6968971; PIDN:CB73682.1; PID:G696911  
 A;Experimental source: serotype O2, strain NCTC 11168  
 C;Genetics:  
 A;Gene: rplX; Cj1696c  
 C;Superfamily: Escherichia coli ribosomal protein L24

Query Match 100.0%; Score 21; DB 2; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 9 KGDS 12

RESULT 9  
 H90114  
 putative small nuclear ribonucleoprotein E [imported] - Guillardia theta nucleomorph  
 C;Species: nucleomorph Guillardia theta  
 A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 C;Accession: H90114  
 R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Perry, S.; Deng, L.T.; Wu, X.; Re  
 Nature 410, 1091-1096, 2001  
 A;Title: The highly reduced genome of an enslaved algal nucleus.  
 A;Reference number: A99082; MUID:11323671; PMID:11323671

Query Match 100.0%; Score 21; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 66 KGDS 69

RESULT 10  
 I55584  
 complement C4A - human (fragment)  
 C;Species: Homo sapiens (man)  
 C;Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jul-2000

Query Match 100.0%; Score 21; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 66 KGDS 69

RESULT 11  
 E69972  
 conserved hypothetical protein yrbF - Bacillus subtilis  
 C;Species: Bacillus subtilis  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C;Accession: E69972  
 R;Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Altoni, G.; Azevedo, V.; Bertei  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galen  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, F  
 A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Yanchin, A.; Yano, K.; Yoshida, F  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377

Query Match 100.0%; Score 21; DB 2; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 43 KGDS 46

RESULT 12  
 D82712  
 histone-like protein Xfl190 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 03-Nov-2000  
 C;Accession: D82712  
 R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Query Match 100.0%; Score 21; DB 2; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 |||||  
 DB 43 KGDS 46

Fri Apr 16 11:04:15 2004

Proc. Natl. Acad. Sci. U.S.A. 75, 615-619, 1978

A:Title: Nucleotide sequence of an insertion element, IS1.

A:Reference number: A93826; MUID:78137003; PMID:273224

A:Accession: D93826

A:Molecule type: DNA

A:Residues: 1-96 <OHT>

R:Johnsrud, L.

Mol. Gen. Genet. 169, 213-218, 1979

A:Title: DNA sequence of the transposable element IS1.

A:Reference number: A93117; MUID:79177885; PMID:375010

A:Accession: D93117

A:Molecule type: DNA

A:Residues: 1-91, 'R', '93-96 <OHT>

A:Cross-references: GB:D10483; GB:J01597; GB:J01683; GB:J01706; GB:K01298; GB:K01990; G

A:Experimental source: strain K-12

R:Kura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizob

submitted to the EMBL Data Library, December 1992

A:Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-

A:Reference number: S40531

A:Accession: S40544

A:Molecule type: DNA

A:Residues: 1-91, 'R', '93-96 <YUR>

A:Cross-references: EMBL:D10483; NID:g216434; PIDN:BAA01299.1; PID:g216448

A:Experimental source: strain K-12

C:Genetics:

A:Mobile element: insertion sequence IS1

C:Superfamily: Escherichia coli insertion sequence IS1 hypothetical 11K protein

Query Match 100.0%; Score 21; DB 1; Length 96;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

Qy 1 KGDS 4

Db 65 KGDS 68

# RESULT 15

I40616

DNA-binding protein Cj0913c [similarity] - Campylobacter jejuni (strain NCTC 11168)

N:Alternate names: HcJ protein; hupB protein

C:Species: Campylobacter jejuni

C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 03-Jun-2002

C:Accession: I40616; H81364

R:Konkel, M.E.; Marconi, R.T.; Mead, D.J.; Cieplak, W.

Gene 146, 83-86, 1994

A:Title: Cloning and expression of the hup encoding a histone-like protein of Campyloba

A:Reference number: I40616; MUID:94341574; PMID:8063109

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-98 <RES>

A:Cross-references: GB:I25627; NID:9409576; PIDN:AAA60954.1; PID:g409577

A:Experimental source: isolate F38011

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81364

A:Molecule type: DNA

A:Residues: 1-98 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73170.1; PID:g69683

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: hupB; Cj0913c

C:Superfamily: bacterial DNA-binding protein

C:Keywords: DNA binding; DNA condensation

Query Match 100.0%; Score 21; DB 2; Length 98;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4

Db 10 KGDS 13

# RESULT 14

IERCE9

hypothetical protein, 11K - Escherichia coli insertion sequence IS1

C:Species: Escherichia coli

C>Date: 24-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 16-Jun-2000

C:Accession: D93826; D93117; S40544; A04452

R:Ohtsubo, H.; Ohtsubo, E.

Nature 406, 151-157, 2003

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: D82712

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-94 <SLX>

A:Cross-references: GB:AE003953; GB:AE003849; NID:g9106156; PIDN:AAF84000.1; GSPDB:GN002

A:Experimental source: strain 945c

R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Buenc, M.R.P.; Camargo, A.A.; Camargo, L.B.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.C. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1190

C:Superfamily: bacterial DNA-binding protein

Query Match 100.0%; Score 21; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4

Db 38 KGDS 41

# RESULT 13

T00965

hypothetical protein F20D22.15 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Oct-1999

C:Accession: T00965

R:Vysotskaia, V.S.; Osborne, B.L.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji,

K.; Feng, J.; Kim, C.; Kurtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.;

submitted to the EMBL Data Library, May 1998

A:Description: Arabidopsis thaliana chromosome 1 BAC F20D22 complete sequence.

A:Reference number: Z4214

A:Accession: T00965

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-94 <VYS>

A:Cross-references: EMBL:AC002411; NID:g2570223; PID:g3142296; GSPDB:GN000059; ATSP:F20D2

C:Genetics:

A:Gene: ATSP.F20D22.15

A:Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4

Db 10 KGDS 13

us-09-991-588b-6.rpr

Fri Apr 16 11:04:15 2004

Qy 1 KGDS 4  
|||  
Db 38 KGDS 41

Search completed: April 16, 2004, 07:29:21  
Job time : 10.4 secs



O88472 mus musculus  
P33494 turkey thin  
P15064 dictyosteli  
Q91x40 neisseria m  
Q9klgl neisseria m  
C67768 aquifex aeo  
Q7vmf2 haemophilus  
P43808 haemophilus  
P19747 capripoxvir  
Q8tlb3 methanosarc  
Q8x8x6 escherichia  
Q8f106 escherichia

34 21 100.0 185 1 TR17 MOUSE  
35 21 100.0 186 1 VMA2 TRTV  
36 21 100.0 189 1 RASG\_DICDI  
37 21 100.0 194 1 RNH2 NEIMA  
38 21 100.0 194 1 RNH2 NEIMA  
39 21 100.0 196 1 RNH2 AQUAE  
40 21 100.0 197 1 RNH2 HADU  
41 21 100.0 197 1 RNH2 HABIN  
42 21 100.0 197 1 VHR2\_CAPVK  
43 21 100.0 197 1 YV25\_METAC  
44 21 100.0 198 1 RNH2 EC057  
45 21 100.0 198 1 RNH2\_EC016

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OX protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time 7 seconds  
(without alignments)

29.754 Million cell updates/sec

Title: US-09-991-588B-6

Perfect score: 21

Sequence: 1 KGDS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	21	100.0	65	1 VG51_BPMD2	O64241 mycobacteri
2	21	100.0	76	1 15E1_HUMAN	O43715 homo sapien
3	21	100.0	89	1 YRBF_BACSU	O32052 bacillus su
4	21	100.0	92	1 DBXA_BURPS	O9zf89 burkholderi
5	21	100.0	94	1 DBH_XVLEA	O9pe38 xylella fas
6	21	100.0	96	1 EUTM_SALTY	P41791 salmonella
7	21	100.0	97	1 EUTM_EC01	P77606 escherichia
8	21	100.0	98	1 DBH_CAMJE	Q46121 campylobact
9	21	100.0	101	1 APC2_CANFA	P12278 canis fami
10	21	100.0	111	1 RUBR_ANAVA	O9xb18 anabaena va
11	21	100.0	120	1 GCSH_RHIME	O92q10 rhizobium m
12	21	100.0	124	1 YPER_BACSU	P54164 bacillus su
13	21	100.0	124	1 PRO2_PHYPO	P18322 physearum po
14	21	100.0	125	1 GCSH_LISIN	Q92813 listeria in
15	21	100.0	125	1 GCSH_LISMO	Q9y412 listeria mo
16	21	100.0	133	1 REV_CAEV	P31628 caprine art
17	21	100.0	149	1 RL13_CHLPN	O9z8t7 chlamydia p
18	21	100.0	152	1 GLBP_CHITH	P11582 chironomus
19	21	100.0	153	1 CXSA_YEAST	P00424 saccharomyc
20	21	100.0	153	1 SODC_HYLLA	Q8hxq3 hylobates l
21	21	100.0	153	1 SODC_MACFA	Q8hxq1 macaca fusc
22	21	100.0	153	1 SODC_MACFU	Q8hxq2 macaca fusc
23	21	100.0	153	1 SODC_MACMU	Q8hxq0 macaca mula
24	21	100.0	154	1 YG39_METUA	Q59033 methanococc
25	21	100.0	158	1 WH21_MYXVL	P28850 myxoma viru
26	21	100.0	159	1 RSBW_STAAM	P95843 staphylococ
27	21	100.0	159	1 RSBW_STAAM	Q8rv15 staphylococ
28	21	100.0	160	1 FMA2_BACNO	P17824 bacteroides
29	21	100.0	161	1 YIAL_BACSU	O07636 bacillus su
30	21	100.0	163	1 IFJL_ANASP	Q8yuh8 anabaena sp
31	21	100.0	171	1 YW87_PSEAE	Q8hyv6 pseudomonas
32	21	100.0	179	1 170K_TRVES	P05079 tobacco rat
33	21	100.0	181	1 RNH2_BORBU	O51075 borrelia bu

## ALIGNMENTS

### RESULT 1

VG51\_BPMD2 STANDARD; PRT; 65 AA.  
ID VG51\_BPMD2  
AC O64241;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Gene 51 protein (GP51).  
GN 51.  
OS Mycobacteriophage D29.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=28369;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98300335; PubMed=9636706;  
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;  
RT "Genome structure of mycobacteriophage D29: implications for phage evolution.";  
RL J. Mol. Biol. 279:143-164(1998).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF022214; AAC18491.1; -  
DR PIR; H72805; H72805.  
SQ SEQUENCE 65 AA; 7364 MW; 7051EC4996BF14FF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 65;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4

DB 48 KGDS 51

### RESULT 2

15E1\_HUMAN STANDARD; PRT; 76 AA.  
ID 15E1\_HUMAN  
AC O43715;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein 15E1.1 (Protein HSPC132).  
GN 15E1.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Murphy L.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=20493367; PubMed=11042152;  
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,  
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han X.-G., Zhang J.-W.,  
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.,  
 RA "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells."  
 RN Genome Res. 10:1546-1560(2000).  
 RP (3)  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SIMILARITY: BELONGS TO THE UPF0203 (15E1.1) FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; AL021546; CAB16495.1; -;  
 CC EMBL; AF161481; AAF29096.1; -;  
 CC EMBL; BC002638; AAH02638.1; -;  
 CC PIR; T09476; T09476.  
 CC InterPro; IPR007918; UPF0203.  
 CC Pfam; PF05254; UPF0203; 1.  
 CC SEQUENCE 76 AA; 8786 MW; 00B41AC399D76590 CRC64;  
 CC  
 CC Query Match 100.0%; Score 21; DB 1; Length 76;  
 CC Best Local Similarity 100.0%; Pred. No. 91;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 KGDS 4  
 CC DB 29 KGDS 32  
 CC  
 CC RESULT 3  
 CC YRBF BACSU STANDARD; PRT; 89 AA.  
 CC AC 032052;  
 CC 30-MAY-2000 (Rel. 39, Created)  
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE Hypothetical protein yrbF.  
 CC GN YRBF OR BSJ27700.  
 CC  
 CC Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=1423;  
 CC (1)  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=168;  
 CC TOSATO V., Bolotin A., Bertani I., Valentini I., Bruschi C.V.;  
 CC "A 17.8 kb segment in the spoVB-nacC region of the Bacillus subtilis  
 CC 168 chromosome: sequencing and ruv operon identification."  
 CC Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC (2)  
 CC SEQUENCES FROM N.A.  
 CC STRAIN=168;  
 CC MEDLINE=98044033; PubMed=9384377;  
 CC Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 CC Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 CC Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 CC Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 CC Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 CC Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 CC Entian K.D., Errington J., Fabret C., Ferrali E., Foulger D.,  
 CC Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 CC Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 CC Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 CC Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 CC Joris B., Karata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 CC Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 CC Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 CC Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 CC Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,  
 CC Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 CC Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 CC Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 CC Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 CC Sato T., Scanlon E., Schleich S., Schroeter R., Scoffone F.,  
 CC Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 CC Sorokin A., Tacconi E., Takagi T., Tanaka T., Terpetra P., Tognoni A.,  
 CC Takachi M., Tanakoshi A., Tandenbol M., Vannier F., Vassarotti A.,  
 CC TOSATO V., Uchiyama S., Vandenbol M., Wedler H., Weitzenecker T.,  
 CC Viari A., Wambutt R., Wedler E., Wedler K., Yasumoto K., Yata K.,  
 CC Winters P., Wipat A., Yamamoto H., Yamane K., Yamamoto K., Yatchin A.;  
 CC Yoshida K., Yoshikawa H.F., Zumsteg H., Yoshikawa H., Danchin A.;  
 CC "The complete genome sequence of the Gram-positive bacterium Bacillus  
 CC subtilis."  
 CC Nature 390:249-256(1997).  
 CC -!- SIMILARITY: BELONGS TO THE UPF0092 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; Y15896; CAB75334.1; -;  
 CC EMBL; Z99118; CAB14730.1; -;  
 CC PIR; E69972; E69972.  
 CC Subtilin; BGL3785; yrbF.  
 CC InterPro; IPR003849; yrbF.  
 CC Pfam; PF02699; yrbF; 1.  
 CC TIGRfams; TIGR00739; yrbF; 1.  
 CC Hypothetical protein; Transmembrane; Complete proteome.  
 CC TRANSMEM 4 24 POTENTIAL.  
 CC SEQUENCE 89 AA; 9904 MW; 23F8DED140D7ACAE CRC64;  
 CC  
 CC Query Match 100.0%; Score 21; DB 1; Length 89;  
 CC Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 KGDS 4  
 CC DB 43 KGDS 46  
 CC

```

RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro M.R., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Pacinca A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.J.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Histone-like DNA-binding protein which is capable of
CC wrapping DNA to stabilize it, and thus to prevent its denaturation
CC under extreme environmental conditions (By similarity).
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB003953; AAF84000.1; -.
CC FIC: D82712; D82712.
CC HSSP: P36206; 1B9Z.
CC InterPro: IPR000119; Bac_DNAbind.
CC Pfam: PF00216; Bac DNA binding; 1.
CC PRINTS: PR01727; DNABINDINGHU.
CC ProDom: PD000945; Bac DNAbind; 1.
CC SMART: SM00411; BHL; 1.
CC PROSITE: PS00045; HISTONE LIKE; FALSE NEG.
KW DNA-binding; DNA condensation; Complete proteome.
SQ SEQUENCE 94 AA; 9766 MW; FFDAC9E1E5125399 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 38 KGDS 41

RESULT 6
EUTM_SALTY STANDARD; PRT; 96 AA.
AC P41791;
DT 01-NOV-1995 (Rel. 32, Created);
DT 01-NOV-1995 (Rel. 32, Last sequence update);
DT 10-OCT-2003 (Rel. 42, Last annotation update);
DE Ethanolamine utilization protein eutM precursor.
GN EUTM OR CCHA OR STM245.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

CC Enterobacteriaceae; Salmonella.  
 RX NCBI\_TaxID=602;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 14028;  
 RX MEDLINE=95173114; PubMed=7868611;  
 RA Stojiljkovic I., Baeumler A.J., Heffron F.;  
 RT "Ethanolamine utilization in Salmonella typhimurium: nucleotide  
 RT sequence, protein expression, and mutational analysis of the ccha  
 RT cchB eutE eutG eutH gene cluster.";  
 RL J. Bacteriol. 177:1357-1366(1995).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=LT2;  
 RC MEDLINE=99395039; PubMed=10464203;  
 RA Kofoid E.C., Rappleye C.A., Stojiljkovic I., Roth J.R.;  
 RT "The 17-gene ethanolamine (eut) operon of Salmonella typhimurium  
 RT encodes five homologues of carboxysome shell proteins.";  
 RL J. Bacteriol. 181:5317-5329(1999).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RC MEDLINE=21534948; PubMed=11677609;  
 RA McClelland L., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 CC -!- FUNCTION: May be involved in the formation of a specific  
 CC microcompartment in the cell in which the metabolism of  
 CC potentially toxic by-products takes place.  
 CC -!- PATHWAY: Ethanolamine utilization.  
 CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U18560; A2480207.1; --  
 CC EMBL; AF093749; AAC78116.1; --  
 CC EMBL; AF008811; AAL21359.1; --  
 CC StyGene; SG10551; eutm.  
 CC InterPro; IPR000249; Bact\_microcomp.  
 CC Pfam; PF00936; Bact\_microcomp; 1.  
 CC ProDom; PD003442; Bact\_microcomp; 1.  
 CC PROSITE; PS01139; BACT\_MICROCOMP; 1.  
 CC SIGNAL; Complete proteome.  
 CC SIGNAL 1 25  
 CC CHAIN 26 96  
 CC SEQUENCE 96 AA; 9842 MW; 01710D8AB8BCF6BC CRC64;  
 CC -----  
 CC Query Match 100.08; Score 21; DB 1; Length 96;  
 CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 KGDS 4  
 CC DB 91 KGDS 94  
 CC -----  
 CC RESULT 7  
 CC EUTM\_ECOLI STANDARD; PRT; 97 AA.  
 CC AC P77606;  
 CC DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ethanolamine utilization protein eutm precursor.  
 GN EUTM OR CCHA OR B2457 OR C2982.  
 OS Escherichia coli, and  
 OS Escherichia coli O6.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562, 217992;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RC MEDLINE=97349980; PubMed=9205837;  
 RA Yanamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,  
 RA Oshima T., Oyama S., Saito N., Sampaio G., Satoh Y., Sivasubram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yanagata S., Horiuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli  
 RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O6.H1 / CFT073 / ATCC 700928;  
 RX MEDLINE=22398234; PubMed=12471157;  
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 RT of uropathogenic Escherichia coli.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
 CC -!- FUNCTION: May be involved in the formation of a specific  
 CC microcompartment in the cell in which the metabolism of  
 CC potentially toxic by-products takes place.  
 CC -!- PATHWAY: Ethanolamine utilization.  
 CC -!- SIMILARITY: Belongs to the bacterial microcompartments protein  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE000332; AAC75510.1; ALT\_INIT.  
 CC EMBL; D90874; BAA16335.1; ALT\_INIT.  
 CC EMBL; AE016764; AAN81432.1; ALT\_INIT.  
 CC EcoGene; EG14187; eutm.  
 CC InterPro; IPR000249; Bact\_microcomp.  
 CC Pfam; PF00936; Bact\_microcomp; 1.  
 CC ProDom; PD003442; Bact\_microcomp; 1.  
 CC PROSITE; PS01139; BACT\_MICROCOMP; 1.  
 CC SIGNAL; Complete proteome.  
 CC SIGNAL 1 25  
 CC CHAIN 26 97  
 CC SEQUENCE 97 AA; 9865 MW; 0A4170FD8AB8BCF8 CRC64;  
 CC -----  
 CC Query Match 100.0%; Score 21; DB 1; Length 97;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0;

QY 1 KGDS 4  
DB 91 KGDS 94

RESULT 8  
DBH CAMJE  
ID DBH CAMJE STANDARD; PRT; 98 AA.  
AC Q46121;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE DNA-binding protein HU [HCJ].  
GN HUP OR HUPB OR CJO913C.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Campylobacteraceae; Campylobacter.  
OX NCBI\_TaxID=297;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.  
RC STRAIN=P38011;  
RX MEDLINE=94341574; PubMed=8053109;  
RA Konkel M.E., Marconi R.R., Mead D.J., Cieplak W. Jr.;  
RT "Cloning and expression of the hup gene encoding a histone-like  
protein of Campylobacter jejuni.";  
RL Gene 146:83-86(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Parkham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
CC -!- FUNCTION: Histone-like DNA-binding protein which is capable of  
wrapping DNA to stabilize it, and thus to prevent its denaturation  
under extreme environmental conditions.  
CC -!- SUBUNIT: Homodimer (by similarity).  
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; J25627; AAA60954.1; --  
EMBL; AL139376; CAB73170.1; --  
PIR; I40616; I40616.  
DR HSP; P36206; I88Z.  
DR InterPro; IPR000119; Bac DNAbind.  
DR Pfam; PF00216; Bac DNA binding; 1.  
DR PRINTS; PR01727; DNABINDINGHU.  
DR ProDom; PDC00945; Bac DNAbind; 1.  
DR SMART; SMO0411; BH2; 1.  
DR PROSITE; PS00045; HISTONE LIKE; FALSE NEG.  
KW DNA-binding; DNA condensation; Complete proteome.  
SQ SEQUENCE 98 AA; 10274 MW; E3CC1FE3BEA70333 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4

Db 38 KGDS 41

RESULT 9  
APC2 CANFA  
ID APC2 CANFA STANDARD; PRT; 101 AA.  
AC P12278;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Apolipoprotein C-II precursor (Apo-CII).  
GN APOC2.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87280118; PubMed=2956259;  
RA Datta S., Li W.-H., Ghosh I., Luo C.-C., Chan L.;  
RT "Structure and expression of dog apolipoprotein C-II and C-III mRNAs.  
Implications for the evolution and functional constraints of  
apolipoprotein structure.";  
RL J. Biol. Chem. 262:10588-10593(1987).  
CC -!- FUNCTION: Component of the very low density lipoprotein (VLDL)  
fraction in plasma, and is an activator of several triacylglycerol  
lipases. The association of APOC2 with plasma chylomicrons,  
VLDL, and HDL is reversible, a function of the secretion and  
catabolism of triglyceride-rich lipoproteins, and changes rapidly.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Secreted in plasma.  
CC -!- SIMILARITY: Belongs to the apolipoprotein C2 family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; M17177; AAA30829.1; --  
PIR; A28487; A28487.  
DR InterPro; IPR008019; Apo-CII.  
DR Pfam; PF05355; Apo-CII; 1.  
KW Chylomicron; VLDL; Transport; Lipid transport; Lipid degradation;  
KW Signal.  
FT SIGNAL. 1 22 APOLIPOPROTEIN C-II.  
FT CHAIN 23 101 LIPID.  
FT BINDING 43 51  
FT DOMAIN 55 78 LIPOPROTEIN LIPASE COFACTOR.  
SQ SEQUENCE 101 AA; 11227 MW; 2FFC5C0B9AA447FC CRC64;

Query Match 100.0%; Score 21; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
DB 98 KGDS 101

RESULT 10  
RUBR ANAVA  
ID RUBR ANAVA STANDARD; PRT; 111 AA.  
AC Q9XEL8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Rubredoxin (Rd).  
GN RUE OR HOXR.  
OS Anabaena variabilis.

```

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=11172;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PC9 7937 / ATCC 29413;
RC Happe T., Schiefer W., Boehme H.;
RT "Isolation and characterisation of the ndhCKJ gene-cluster of Anabaena
RT variabilis";
RL Submitted (Oct-1998); to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Rubredoxin is a small nonheme, iron protein lacking
CC acid-labile sulfide. Its single Fe, chelated to 4 Cys, functions
CC as an electron acceptor and may also stabilize the conformation of
CC the molecule. Could be involved in hydrogenase-linked redox
CC processes (By similarity).
CC -!- COFACTOR: Binds 1 iron(III) ion per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the rubredoxin family.
CC -!- SIMILARITY: Contains 1 rubredoxin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ012181; CAB45645.1; -.
CC HSPSP; P00270; IRDG.
CC InterPro; IPR004039; Rubredox.
CC InterPro; IPR001052; Rubredoxin.
CC Pfam; PF00301; rubredoxin; 1.
CC PRINTS; PR00163; RUBREDOXIN.
CC PRODOM; PD001610; Rubredoxin; 1.
CC PROSITE; PS00202; RUBREDOXIN; FALSE NEG.
CC PROSITE; PS0903; RUBREDOXIN-LIKE; 1.
CC Rubredoxin; Electron transport; Iron; Metal-binding.
FT DOMAIN 11 62 RUBREDOXIN-LIKE.
FT METAL 16 16 IRON (BY SIMILARITY).
FT METAL 19 19 IRON (BY SIMILARITY).
FT METAL 49 49 IRON (BY SIMILARITY).
FT METAL 52 52 IRON (BY SIMILARITY).
SQ SEQUENCE 111 AA; 12093 MW; 0B1CEPABH764CB2B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

Qy 1 KGDS 4
Db 27 KGDS 30

RESULT 11
GCSH RHIME STANDARD; PRT; 120 AA.
AC Q92Q10.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glycine cleavage system H protein.
GN GCVH OR R01548 OR SMC02048.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,

Renard C., Thebault P., Vandenbol M., Weidner S., Galibert P.;
"Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021".
Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
-!- FUNCTION: The glycine cleavage system catalyzes the degradation of
glycine. The H protein shuttles the methylamine group of glycine
from the P protein to the T protein (By similarity).
-!- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
(BY similarity).
-!- SUBUNIT: The glycine cleavage system is composed of four proteins:
P, T, L and H (By similarity).
-!- SIMILARITY: Belongs to the gcvh family.
-!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL591787; CAC46127.1; -.
CC HAMAP; MF 00272; -. 1.
CC InterPro; IPR002930; GCV H.
CC InterPro; IPR003016; Lipoyl_BS.
CC Pfam; PF01597; GCV_H; 1.
CC TIGRFAMs; TIGR00527; GCVH; 1.
CC PROSITE; PS00189; LIPOYL; 1.
CC Lipoyl; Complete proteome.
FT BINDING 58 58 LIPOYL (BY SIMILARITY).
SQ SEQUENCE 120 AA; 12727 MW; 5741C4FFED82B437 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

Qy 1 KGDS 4
Db 47 KGDS 50

RESULT 12
YPEP BACSU
ID YPEP BACSU STANDARD; PRT; 120 AA.
AC P54164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ypep.
GN YPEP OR BSU21970.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RA "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kdg loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

```

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kaashara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumaro M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigic C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetlelle D., Porwollik S., Prescott A.M.,  
 RA Prescan E., Puji P., Purnelle B., Rapoport S., Rev M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,  
 RA Sato T., Scania E., Schleich S., Schroeter R., Scorfone P.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi K., Takemaru K.,  
 RA Takeuchi X., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 RA Tosato V., Uchiyara S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L77246; AAA96620.1; -;  
 DR EMBL; Z99115; CAB14115.1; -;  
 DR PIR; F69934; F69934.  
 DR Subtilist; BG1609; ypeP.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 120 AA; 13720 MW; D3P4FFA765E0A867 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 DB 40 KGDS 43  
 RESULT 13  
 PRO2\_PHYPO STANDARD; PRT; 124 AA.  
 ID PRO2\_PHYPO STANDARD; PRT; 124 AA.  
 AC P18322;  
 DT 01-NOV-1992 (Rel. 16, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Profilin P.  
 GN PROF.  
 OS Physarum polycephalum (Slime mold).  
 OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;  
 OC Physarum.  
 OX NCBI\_TaxID=5791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90321467; PubMed=2372376;  
 RA Binette F., Benard M., Laroche A., Pierron G., Lemieux G.,  
 RA Pallotta D.;  
 RT "Cell-specific expression of a profilin gene family";  
 RL DNA Cell Biol. 9:323-334(1990).  
 RP SEQUENCE.  
 RX MEDLINE=91096174; PubMed=2209623;

RA Takagi T., Mabuchi I., Hosoya H., Furuhashi K., Hatano S.;  
 RT "Primary structure of profilins from two species of Echinoidea and  
 RT Physarum polycephalum.";  
 RL Eur. J. Biochem. 192:777-781(1990).  
 CC -!- FUNCTION: Binds to actin and affects the structure of the  
 CC cytoskeleton. At high concentrations, profilin prevents the  
 CC polymerization of actin, whereas it enhances it at low  
 CC concentrations. By binding to PIP2, it inhibits the formation of  
 CC IF3 and DG.  
 CC -!- SUBUNIT: Occurs in many kinds of cells as a complex with monomeric  
 CC actin in a 1:1 ratio.  
 CC -!- SIMILARITY: Belongs to the profilin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M36038; AAD13630.1; -;  
 DR PIR; B35273; B35273.  
 DR PIR; S13199; S13199.  
 DR HSP; P25816; LCOA.  
 DR InterPro; IPR002097; Profilin.  
 DR InterPro; IPR003455; Profilin\_plant.  
 DR Pfam; PF00235; profilin; 1.  
 DR PRINTS; PR00392; PROFILIN.  
 DR PRINTS; PR01640; PROFILINPLNT.  
 DR SMART; SM00392; PROF; 1.  
 DR PROSITE; PS00414; PROFILIN; 1.  
 KW Actin-binding; Cytoskeleton; Multigene family; Acetylation.  
 FT INIT MET 0 0  
 FT MOD RES 1 1 ACETYLATION.  
 FT CONFLICT 7 7 D -> H (IN REF. 1).  
 FT CONFLICT 85 85 V -> S (IN REF. 2).  
 SQ SEQUENCE 124 AA; 13061 MW; F2343F75EA724F6 CRC64;  
 Query Match 100.3%; Score 21; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 DB 70 KGDS 73  
 RESULT 14  
 GCSH\_LISIN STANDARD; PRT; 125 AA.  
 ID GCSH\_LISIN STANDARD; PRT; 125 AA.  
 AC Q92813;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Glycine cleavage system H protein.  
 GN GCVH OR L1N2519.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Clp 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Cherouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf C., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

```

RA Remme B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC glycine. The H protein shuttles the methylamine group of glycine
CC from the P protein to the T protein (By similarity).
CC -1- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC (By similarity).
CC -1- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC P, T, L and H (By similarity).
CC -1- SIMILARITY: Belongs to the gcvH family.
CC -1- SIMILARITY: Contains 1 lipoyl-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL596172; CAC97746.1; -
CC PIR; AB1747; AB1747.
CC ListList; LINO2519; -.
CC HAMAP; MF_00272; -.
CC InterPro; IPR002930; GCV_H.
CC InterPro; IPR003016; Lipoyl_BS.
CC Pfam; PF01597; GCV_H; 1.
CC TIGRFAMs; TIGR00527; gcvH; 1.
CC PROSITE; PS00189; LIPOYL; 1.
CC Lipoyl; Complete proteome.
CC BINDING 63 63 LIPOYL (BY SIMILARITY).
CC FT BINDING 125 AA; 13732 MW; 94F00032FBCFA152 CRC64;
CC SEQUENCE 125 AA; 13732 MW; 134EE1E37C8980B5 CRC64;
CC
CC Query Match 100.0%; Score 21; DB 1; Length 125;
CC Best Local Similarity 100.0%; Pred. No. 1.6e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 KGDS 4
CC Db 52 KGDS 55
CC
CC RESULT 15
CC GCSH LISMO STANDARD; PRT; 125 AA.
CC AC Q8V4L2;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Glycine cleavage system H protein.
CC GN GCVH OR LMO2425.
CC OS Listeria monocytogenes.
CC CC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CC NCBI_TaxID=1639;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=EGD-e / Serovar 1/2a;
CC RX MEDLINE=21537279; PubMed=11679569;
CC RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
CC RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
CC RA Charbit A., Chetoui F., Couve E., Ge Baruv A., Denoux P.,
CC RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
CC RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
CC RA Gautier L., Goebel W., Gomez-Lopez N., Hain J., Jackson D.,
CC RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunz F., Nedjari H.,
CC RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
CC RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
CC RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
CC RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
CC FT "Comparative genomics of Listeria species.";
CC RL Science 294:849-852(2001).

```

Search completed: April 16, 2004, 07:22:13  
Job time : 9 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 31.6 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-6  
Perfect score: 21  
Sequence: 1 KGDS 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 3  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	33	5 Q7YV06	Q7YV06 trypanosoma
2	21	100.0	43	10 O64386	O64386 solanum tub
3	21	100.0	45	16 Q8F5X8	Q8F5X8 leptospira
4	21	100.0	46	6 Q9TIG4	Q9TIG4 ateleles belz
5	21	100.0	53	16 Q928W3	Q928W3 listeria in
6	21	100.0	54	11 Q8K4F9	Q8K4F9 rattus norv
7	21	100.0	55	7 Q8WLW0	Q8WLW0 ginglymosto
8	21	100.0	58	2 Q9L5Q8	Q9L5Q8 salmonella
9	21	100.0	59	4 Q9UD54	Q9UD54 homo sapien
10	21	100.0	59	4 Q9UD53	Q9UD53 homo sapien
11	21	100.0	59	16 Q8X373	Q8X373 escherichia
12	21	100.0	62	10 Q84L52	Q84L52 hordeum vul
13	21	100.0	63	16 Q8VIY8	Q8VIY8 brucella me
14	21	100.0	63	16 Q8FYW8	Q8FYW8 brucella su
15	21	100.0	63	16 Q82TC3	Q82TC3 nitrosomona
16	21	100.0	66	4 Q9HIA1	Q9HIA1 homo sapien

17	21	100.0	57	2 Q93PY7	Q93PY7 pseudomonas
18	21	100.0	69	2 Q7WTS2	Q7WTS2 paracoccus
19	20	100.0	70	2 Q9EXE4	Q9EXE4 lactobacilli
20	21	100.0	71	11 Q9BR50	Q9BR50 mus musculu
21	21	100.0	72	12 Q91RY9	Q91RY9 lettuce mos
22	21	100.0	72	12 Q91RZ1	Q91RZ1 lettuce mos
23	21	100.0	72	12 Q91RY6	Q91RY6 lettuce mos
24	21	100.0	73	10 Q7AB59	Q7AB59 oryza sativ
25	21	100.0	73	12 Q9WMF3	Q9WMF3 tt virus. h
26	21	100.0	73	16 Q8EB86	Q8EB86 shewanella
27	21	100.0	74	12 Q91Q86	Q91Q86 tt virus. c
28	21	100.0	74	12 Q9YZJ5	Q9YZJ5 tt virus. o
29	21	100.0	74	12 Q91Q85	Q91Q85 tt virus. o
30	21	100.0	74	12 Q91Q84	Q91Q84 tt virus. o
31	21	100.0	75	16 Q8X293	Q8X293 escherichia
32	21	100.0	75	16 Q8PKJ6	Q8PKJ6 xanthomonas
33	21	100.0	76	16 Q7ULR4	Q7ULR4 rhodospirell
34	21	100.0	77	16 Q9PLY2	Q9PLY2 campylobact
35	21	100.0	78	10 Q9AVY1	Q9AVY1 guillardia
36	21	100.0	78	10 Q9SWZ2	Q9SWZ2 lycopersico
37	21	100.0	79	7 Q30211	Q30211 homo sapien
38	21	100.0	83	12 Q9J7M7	Q9J7M7 tt virus. o
39	21	100.0	85	10 Q8LJ84	Q8LJ84 oryza sativ
40	21	100.0	85	12 Q919Y6	Q919Y6 tt virus. o
41	21	100.0	85	16 Q8ABW8	Q8ABW8 bacteroides
42	21	100.0	86	10 Q7XV85	Q7XV85 oryza sativ
43	21	100.0	86	12 Q9WQD9	Q9WQD9 tt virus. h
44	21	100.0	89	1 O06113	O06113 methanococc
45	21	100.0	89	10 Q94JJ1	Q94JJ1 oryza sativ

## ALIGNMENTS

RESULT 1  
Q7YV06  
ID Q7YV06 PRE-LIMINARY; PRT; 33 AA.  
AC QYV06; 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN TB927.2.4270.  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GUTat10.1;  
RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,  
RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,  
RA Blean N., Khalak H.G., Lin X., Mason T., Hannick L., Caler E.,  
RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pai G.,  
RA Van Aken S., Utterback T., Haas B., Koo H.L., Umayam L., Suh B.,  
RA Gerrard C., Leach V., Qi R., Zhou S., Schwartz D., Feldblyum T.,  
RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,  
RA Adams M.D., Fraser C.M., Donelson J.E.;  
FT "The sequence and analysis of Trypanosoma brucei chromosome II.";  
RL Nucleic Acids Res. 0:0-0(2003).  
DR EMBL; AE017169; AAQ15884.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 33 AA; 3812 MW; 18931C9AESABF41F CRC64;

Query Match 100.0%; Score 21; DB 5; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4  
Db 17 KGDS 20

RESULT 2

064386 PRELIMINARY; PRT; 43 AA.  
 ID O64386;  
 AC O64386;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Reverse transcriptase (Fragment).  
 OS Solanum tuberosum (potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Desiree; TISSUE=Leaf;  
 RA Kuipers A.G.J., Garriga-Caldere F., Heslop-Harrison J.S., Jacobsen E.;  
 RT "Characterization and physical localization of Ty1-copia-like  
 RT retrotransposons in potato and tomato."  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ228807; CAA13064.1; -;  
 DR PIR; T07124; T07124.  
 FT NON\_TER 1 1  
 FT NON\_TER 43 43  
 SQ SEQUENCE 43 AA; 4996 XM; 798AD74B0142CFE3 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 DB 29 KGDS 32

RESULT 3  
 Q8F5K8 PRELIMINARY; PRT; 45 AA.  
 ID Q8F5K8;  
 AC Q8F5K8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN LA1672.  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB011345; AAM48872.1; -;  
 DR EMBL; AB011345; AAM48872.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 45 AA; 5273 MW; 75ABB66C3CCF1E0A CRC64;

Query Match 100.0%; Score 21; DB 16; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 DB 34 KGDS 37

RESULT 4  
 Q9ITG4 PRELIMINARY; PRT; 45 AA.  
 ID Q9ITG4;  
 AC Q9ITG4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Heart chitinase (Fragment).  
 GN CMA.

OS Ateles belzebuth chamek (Chamek spider monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aelinae; Ateles.  
 OX NCBI\_TaxID=118643;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20169179; PubMed=0702662;  
 RA Canavez F.C., Moreira M.A., Bonvicino C.R., Parham P., Seuanez H.N.;  
 RT "Evolutionary disruptions of human syntenic groups 3, 12, 14, and 15  
 RT in Ateles belzebuth chamek (Platyrrhini, Primates).";  
 RL Cytogenet. Cell Genet. 87:182-188(1999).  
 DR EMBL; AF099174; AAF21848.1; -;  
 DR HSP; P23946; IKLT.  
 DR MEROPS; S01.140; -;  
 DR GO; GO:0008233; F:Peptidase activity; IEA.  
 DR GO; GO:0004295; F:Trypsin activity; IEA.  
 DR GO; GO:0006508; P:Proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR01254; Peptidase\_S1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 1 1  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 5003 MW; 481EA020A265D29F CRC64;

Query Match 100.0%; Score 21; DB 6; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 DB 35 KGDS 38

RESULT 5  
 Q928W3 PRELIMINARY; PRT; 53 AA.  
 ID Q928W3;  
 AC Q928W3;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein lin2419.  
 GN LIN2419.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Denoux P.,  
 RA Donann E., Dominguez-Bernal G., Duchaud E., Durand L., Dussurget C.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-W., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of Listeria species.";  
 RL Science 294:849-852(2001).  
 DR EMBL; AL596172; CAC97646.1; -;  
 DR PIR; AF1734; AF1734.  
 DR Listlist; LIN02419; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 53 AA; 6104 MW; 92F71761F2435DE6 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KGDS 4
DB 35 KGDS 38

RESULT 6
Q8K4F9 PRELIMINARY; PRT; 54 AA.
ID Q8K4F9;
AC Q8K4F9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Fragile-X mental retardation related protein 2 (Fragment).
GN FXR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=BN;
RX MEDLINE=22032986; PubMed=12036299;
RA Behboudi A., Roshani L., Lundin L., Stahl F., Levan K.K., Levan G.;
RT "The functional significance of absence: the chromosomal segment
RT harboring tp53 is absent from the t55 rat radiation hybrid mapping
RT panel.";
RL Genomics 79:844-848(2002).
DR EMBL; AF410814; AAM48519.1; -.
FT NON_TER 1 1
FT NON_TER 54 54
SQ SEQUENCE 54 AA; 5951 MW; C17B7945FBF938DC CRC64;

Query Match 100.0%; Score 21; DB 11; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 43 KGDS 46

RESULT 7
Q8WLW0 PRELIMINARY; PRT; 55 AA.
ID Q8WLW0;
AC Q8WLW0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MHC class I antigen (Fragment).
GN UAA.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]_TaxID=7801;
RP SEQUENCE FROM N.A.
RX MEDLINE=21636493; PubMed=11777971;
RA Ohta Y., McKinney E.C., Cristicciello M.F., Flajnik M.F.;
RT "Proteasome, Transporter Associated with Antigen Processing, and Class
RT I Genes in the Nurse Shark Ginglymostoma cirratum: Evidence for a
RT Stable Class I Region and MHC Haplotype Lineages.";
RL J. Immunol. 168:771-781(2002).
DR EMBL; AF363581; AAL59860.1; -.
FT NON_TER 1 1
FT NON_TER 55 55
SQ SEQUENCE 55 AA; 5617 MW; 3F7C3263C463F948 CRC64;

Query Match 100.0%; Score 21; DB 7; Length 55;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 43 KGDS 46

RESULT 8
Q9L508 PRELIMINARY; PRT; 58 AA.
ID Q9L508;
AC Q9L508;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orf, hypothetical protein.
GN R0073.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]_TaxID=601;
RP SEQUENCE FROM N.A.
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
DR EMBL; AF250878; AAF69911.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 58 AA; 6733 MW; 95D6B5190E6681FF CRC64;

Query Match 100.0%; Score 21; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 7 KGDS 10

RESULT 9
Q9UD54 PRELIMINARY; PRT; 59 AA.
ID Q9UD54;
AC Q9UD54;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coagulation factor VII (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=94318476; PubMed=8043443;
RA Bernardi F., Liney D.J., Patrachini P., Gemmati D., Legnani C.,
RA Arcieri P., Pinotti M., Redaelli R., Ballerini G., Pemberton S.;
RT "Molecular defects in CRM+ factor VII deficiencies: modelling of
RT missense mutations in the catalytic domain of FVII.";
RL Br. J. Haematol. 86:610-618(1994).
DR HSP; P08709; 1FAK.
DR GC; GO:0008233; F:peptidase activity; IEA.
DR GC; GO:0004295; F:trypsin activity; IEA.
DR GC; GO:0008508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR01254; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 59 AA; 6406 MW; 1E334F84031BE1D1 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

SQ SEQUENCE 59 AA; 6795 MW; 5B6B409F62CDBA89 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 59;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
Db 38 KGDS 41

RESULT 10  
Q9UD53 PRELIMINARY; PRT; 59 AA.

AC Q9UD53;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Coagulation factor VII (Fragment).  
OS Homo sapiens (Human).  
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94318476; PubMed=8043443;  
RA Bernardi F., Linney D.L., Patraccini P., Gemmati D., Legnani C.,  
RA Arcieri P., Pinotti M., Redaelli R., Ballerini G., Pemberton S.;  
RT "Molecular defects in CRF+ factor VII deficiencies: modelling of  
RT Missense mutations in the catalytic domain of FVII.";  
RL Br. J. Haematol. 86:610-618(1994).  
DR HSP; P08709; 1PAK.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR Pfam; PF00089; trypsin; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydroxylase; Protease; Serine protease.  
SQ SEQUENCE 59 AA; 6468 MW; 4E374F90525AE6CA CRC64;

Query Match 100.0%; Score 21; DB 4; Length 59;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
Db 48 KGDS 51

RESULT 11  
Q8X373 PRELIMINARY; PRT; 59 AA.

AC Q8X373;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein Ecs2890.  
GN ECS2890.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OC NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Har C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa K., Ogasawara N., Yasunaga T.,  
RA Kohara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
DR EMBL; AF002560; BAB36313.1; --  
DR PIR; B90990; B90990.  
KW Hypothetical protein.

SQ SEQUENCE 59 AA; 6795 MW; 5B6B409F62CDBA89 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 59;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
Db 38 KGDS 41

RESULT 12  
Q84L52 PRELIMINARY; PRT; 62 AA.

AC Q84L52;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Metallothionein-like protein type 1.  
GN MT-1B.  
OS Hordeum vulgare var. distichum (Two-rowed barley).  
OC Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Hordeum.  
OC NCBI\_TaxID=112509;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=cv. Steffi; TISSUE=Primary leaves;  
RA Heise J., Krejci S., Miersch J., Humbeck K.;  
RT "Expression patterns of six novel metallothionein-like genes during  
RT senescence and metal-stress in barley (Hordeum vulgare) primary  
RT leaves.";  
RL Submitted (APR-2003) to the EMBL/GenBank/DOBJ databases.  
DR EMBL; AJ555613; CAD88266.1; --  
SQ SEQUENCE 62 AA; 6657 MW; 494E5F9A380906B1 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 62;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
Db 19 KGDS 22

RESULT 13  
Q8YIY8 PRELIMINARY; PRT; 63 AA.

AC Q8YIY8;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein BMEI0301.  
GN BMEI0301.  
OS Brucella melitensis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Brucellaceae; Brucella.  
OC NCBI\_TaxID=29459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=16M / ATCC 23456 / Biotype 1;  
RX MEDLINE=20020109; PubMed=11756688;  
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,  
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Renik G.,  
RA Jablonski L., Larsen N., D'Souza M., Betnal A., Mazur M., Goltsman E.,  
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
RA Haselkorn R., Kyprides N., Overbeek R.;  
RT "The genome sequence of the facultative intracellular pathogen  
RT Brucella melitensis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
DR EMBL; AF009472; BAL51482.1; --  
DR PIR; AG3289; AG3289.  
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 63 AA; 6954 MW; 848D1F2E39E3F3C1 CRC64;  
 Query Match 100.0%; Score 21; DB 16; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 DB 51 KGDS 54  
 RESULT 14  
 Q8EYWB PRELIMINARY; PRT; 63 AA.  
 AC Q8FYWB;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN BR1739.  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 CX NCBI\_TaxID=29461;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RX MEDLINE=22247741; PubMed=12271122;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
 RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,  
 RA Daugherty S.C., Deboy R.F., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek C., Van Aken S.E.,  
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;  
 RT "The Brucella suis genome reveals fundamental similarities between  
 animal and plant pathogens and symbionts."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
 DR EMBL; AE014465; AAJ30638.1; -.  
 DR PIR; AG3289; AG3289.  
 DR TIGR; BR1739; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 63 AA; 6954 MW; 848D1F2E39E3F3C1 CRC64;  
 Query Match 100.0%; Score 21; DB 16; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 DB 51 KGDS 54  
 RESULT 15  
 Q82TC3 PRELIMINARY; PRT; 63 AA.  
 AC Q82TC3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative transmembrane protein.  
 GN NE1977.  
 OS Nitrosomonas europaea.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
 OC Nitrosomonadaceae; Nitrosomonas.  
 CX NCBI\_TaxID=915;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19718 / IFO 14298;  
 RX MEDLINE=22586410; PubMed=2700255;  
 RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,  
 RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,  
 RA Arciero D.M., Holmes N.G., Whittaker M.M., Arp D.J.;  
 RT "Complete genome sequence of the ammonia-oxidizing bacterium and

RT obligate chemolithoautotroph Nitrosomonas europaea.";  
 RL J. Bacteriol. 185:2759-2773(2003).  
 DR EMBL; BX321863; CAD95888.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 KW Transmembrane; Complete proteome.  
 SQ SEQUENCE 63 AA; 7286 MW; C9696A678A996771 CRC64;  
 Query Match 100.0%; Score 21; DB 16; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 DB 27 KGDS 30  
 Search completed: April 16, 2004, 07:27:47  
 Job time : 34.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:28:08 ; Search time 40.35 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588B-7

Perfect score: 34  
Sequence: 1 GRGDSF 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PTCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	6	9	US-09-837-969A-28
2	34	100.0	6	9	US-09-837-969A-46
3	34	100.0	6	9	US-09-888-260-15
4	34	100.0	6	9	US-09-961-834-1
5	34	100.0	6	9	US-09-364-597A-21
6	34	100.0	6	9	US-09-177-843-1
7	34	100.0	6	9	US-09-841-321A-28
8	34	100.0	6	9	US-09-841-321A-46
9	34	100.0	6	9	US-09-871-974-3
10	34	100.0	6	9	US-09-972-772-31
11	34	100.0	6	11	US-09-991-588B-7
12	34	100.0	6	12	US-09-961-381A-3
13	34	100.0	6	12	US-09-380-682-63
14	34	100.0	6	12	US-10-136-187-42
15	34	100.0	6	13	US-10-001-945-31

16	34	100.0	6	13	US-10-007-270-32	Sequence 12, Appl
17	34	100.0	6	13	US-10-007-270-35	Sequence 35, Appl
18	34	100.0	6	13	US-10-114-176-2	Sequence 2, Appl
19	34	100.0	6	14	US-10-046-801-5	Sequence 5, Appl
20	34	100.0	6	14	US-10-237-850-83	Sequence 83, Appl
21	34	100.0	6	14	US-10-138-935-31	Sequence 31, Appl
22	34	100.0	6	14	US-10-262-435-5	Sequence 5, Appl
23	34	100.0	6	14	US-10-325-021-9	Sequence 9, Appl
24	34	100.0	6	14	US-10-323-013-2	Sequence 2, Appl
25	34	100.0	6	14	US-10-439-532-15	Sequence 15, Appl
26	34	100.0	6	14	US-10-086-208-5	Sequence 5, Appl
27	34	100.0	6	15	US-10-259-609-14	Sequence 14, Appl
28	34	100.0	6	15	US-10-182-432-31	Sequence 3, Appl
29	34	100.0	7	8	US-08-987-756-1	Sequence 1, Appl
30	34	100.0	7	9	US-09-364-597A-22	Sequence 22, Appl
31	34	100.0	7	10	US-09-911-569-8	Sequence 8, Appl
32	34	100.0	7	11	US-09-991-588B-10	Sequence 10, Appl
33	34	100.0	7	14	US-10-200-879-8	Sequence 8, Appl
34	34	100.0	7	14	US-10-420-029-3	Sequence 3, Appl
35	34	100.0	7	15	US-10-297-229-74	Sequence 74, Appl
36	34	100.0	7	16	US-10-208-894A-5	Sequence 5, Appl
37	34	100.0	8	9	US-09-096-749A-96	Sequence 96, Appl
38	34	100.0	8	9	US-09-892-071-3	Sequence 3, Appl
39	34	100.0	8	9	US-09-753-126-6	Sequence 6, Appl
40	34	100.0	8	10	US-09-903-412-96	Sequence 96, Appl
41	34	100.0	8	14	US-10-174-717A-96	Sequence 96, Appl
42	34	100.0	8	14	US-10-165-155-96	Sequence 96, Appl
43	34	100.0	8	14	US-10-190-162-96	Sequence 96, Appl
44	34	100.0	8	15	US-10-330-697-6	Sequence 6, Appl
45	34	100.0	9	9	US-09-892-071-11	Sequence 11, Appl

## ALIGNMENTS

RESULT 1  
US-09-837-969A-28  
; Sequence 28, Application US/09837969A  
; Patent No. US20020038150A1  
; GENERAL INFORMATION:

; APPLICANT: Ury, Dan  
; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
; FILE REFERENCE: BERL-020/03US  
; CURRENT APPLICATION NUMBER: US/09/837,969A  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 09/258,723  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: US 60/087155  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: US 60/076297  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(6)  
; OTHER INFORMATION: Synthetic  
US-09-837-969A-28

Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSF 6  
| | | | |  
Db 1 GRGDSF 6

RESULT 2

US-09-837-969A-46  
; Sequence 46, Application US/09837969A  
; Patent No. US20020038150A1  
; GENERAL INFORMATION:  
; APPLICANT: Urry, Dan  
; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
; FILE REFERENCE: BERL-020/030US  
; CURRENT APPLICATION NUMBER: US/09/837,969A  
; CURRENT FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 09/258,723  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: US 60/087155  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: US 60/076297  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 46  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(6)  
; OTHER INFORMATION: Synthetic  
US-09-837-969A-46  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDSP 6  
DB 1 GRGDSP 6  
|||||  
|  
  
RESULT 3  
US-09-888-260-15  
; Sequence 15, Application US/09888260  
; Patent No. US20020068304A1  
; GENERAL INFORMATION:  
; APPLICANT: Bioplastics Research, Ltd.  
; TITLE OF INVENTION: Bioelastomer Nanomachines and Biosensors  
; FILE REFERENCE: BERL-030/01US  
; CURRENT APPLICATION NUMBER: US/09/888,260  
; CURRENT FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: 60/213,364  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(6)  
; OTHER INFORMATION: Synthetic  
US-09-888-260-15  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDSP 6  
DB 1 GRGDSP 6  
|||||  
|  
  
US-09-837-969A-46  
; Patent No. US20020081726A1  
; GENERAL INFORMATION:  
; APPLICANT: Russell et al.  
; TITLE OF INVENTION: MICROFABRICATION OF MEMBRANES FOR THE GROWTH OF CELLS  
; FILE REFERENCE: 27611/37761  
; CURRENT APPLICATION NUMBER: US/09/961,834  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: US 60/235,094  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Fibronectin ligand receptor  
US-09-961-834-1  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDSP 6  
DB 1 GRGDSP 6  
|||||  
|  
  
US-09-364-597A-21  
; Sequence 21, Application US/09364597A  
; Patent No. US20020103130A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Koivunen, Erkki  
; TITLE OF INVENTION: No. US20020103130A1e1 Integrin-Binding Peptides  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/364,597A  
; FILING DATE: 30-JUL-1999  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/158,001  
; FILING DATE: 24-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,861  
; FILING DATE: 04-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 3419  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (858) 535-9001  
; TELEFAX: (858) 535-8949  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-364-597A-21  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-837-969A-46  
; Sequence 46, Application US/09837969A  
; Patent No. US20020038150A1  
; GENERAL INFORMATION:  
; APPLICANT: Urry, Dan  
; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
; FILE REFERENCE: BERL-020/030US  
; CURRENT APPLICATION NUMBER: US/09/837,969A  
; CURRENT FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 09/258,723  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: US 60/087155  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: US 60/076297  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 46  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(6)  
; OTHER INFORMATION: Synthetic  
US-09-837-969A-46  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDSP 6  
DB 1 GRGDSP 6  
|||||  
|  
  
RESULT 4  
US-09-888-260-15  
; Sequence 15, Application US/09888260  
; Patent No. US20020068304A1  
; GENERAL INFORMATION:  
; APPLICANT: Bioplastics Research, Ltd.  
; TITLE OF INVENTION: Bioelastomer Nanomachines and Biosensors  
; FILE REFERENCE: BERL-030/01US  
; CURRENT APPLICATION NUMBER: US/09/888,260  
; CURRENT FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: 60/213,364  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(6)  
; OTHER INFORMATION: Synthetic  
US-09-888-260-15  
  
Query Match 100.0%; Score 34; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDSP 6  
DB 1 GRGDSP 6  
|||||  
|  
  
US-09-961-834-1  
; Sequence 1, Application US/09961834

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 1 GRGDSP 6

## RESULT 6

US-09-177-843-1  
 ; Sequence 1, Application US/09177843  
 ; Patent No. US20020114804A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE  
 ; APPLICANT: CITY OF NEW YORK  
 ; TITLE OF INVENTION: A METHOD OF PREVENTING AND TREATING  
 ; TITLE OF INVENTION: BACTERIAL INFECTION OF SUTURES AND  
 ; TITLE OF INVENTION: PROSTHETIC DEVICES, AND PROMOTING  
 ; TITLE OF INVENTION: INGRESS OF LEUKOCYTES INTO TUMOR  
 ; TITLE OF INVENTION: FOCI  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/177,843  
 ; FILING DATE: April 22, 1997  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:

NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 48940-A-PCT/JPW/JKM  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-278-0400  
 ; TELEFAX: 212-391-0525  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; OTHER INFORMATION: Synthetic  
 ; US-09-177-843-1

Query Match 100.0%; Score 34; DB 9; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 1 GRGDSP 6

## RESULT 7

US-09-841-321A-28  
 ; Sequence 28, Application US/09841321A  
 ; Patent No. US20020116069A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Urry, Dan  
 ; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
 ; FILE REFERENCE: BERL-020/04US  
 ; CURRENT APPLICATION NUMBER: US/09/841,321A  
 ; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: US 09/258,723  
 ; PRIOR FILING DATE: 1999-02-26  
 ; PRIOR APPLICATION NUMBER: US 60/087155  
 ; PRIOR FILING DATE: 1998-05-29  
 ; PRIOR APPLICATION NUMBER: US 60/076297  
 ; PRIOR FILING DATE: 1998-02-27  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 28  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(6)  
 ; OTHER INFORMATION: Synthetic  
 ; US-09-841-321A-28

Query Match 100.0%; Score 34; DB 9; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 1 GRGDSP 6

## RESULT 8

US-09-841-321A-46  
 ; Sequence 46, Application US/09841321A  
 ; Patent No. US20020116069A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Urry, Dan  
 ; TITLE OF INVENTION: Injectable Implants For Tissue Augmentation and Restoration  
 ; FILE REFERENCE: BERL-020/04US  
 ; CURRENT APPLICATION NUMBER: US/09/841,321A  
 ; CURRENT FILING DATE: 2001-04-30  
 ; PRIOR APPLICATION NUMBER: US 09/258,723  
 ; PRIOR FILING DATE: 1999-02-26  
 ; PRIOR APPLICATION NUMBER: US 60/087155  
 ; PRIOR FILING DATE: 1998-05-29  
 ; PRIOR APPLICATION NUMBER: US 60/076297  
 ; PRIOR FILING DATE: 1998-02-27  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 46  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(6)  
 ; OTHER INFORMATION: Synthetic  
 ; US-09-841-321A-46

Query Match 100.0%; Score 34; DB 9; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 1 GRGDSP 6

## RESULT 9

US-09-871-974-3  
 ; Sequence 3, Application US/09871974  
 ; Patent No. US20020147136A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VON WRONSKI, MATHEW A.  
 ; APPLICANT: MARINELLI, EDMUND R.  
 ; APPLICANT: NUNN, ADRIAN D.  
 ; APPLICANT: PILLAI, RADHAKRISHNA



```

; TITLE OF INVENTION: Compositionand Method for Bone Regeneration
; FILE REFERENCE: 1008-120. US
; CURRENT APPLICATION NUMBER: US/09/991,588B
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 09/122,348
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 7
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Purchased commercially or sequence is synthesized
US-09-991-588B-7

Query Match 100.0%; Score 34; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 3;

QY 1 GRGDSP 6
DB 1 GRGDSP 6

RESULT 12
US-09-961-381A-3
; Sequence 3, Application US/099611381A
; Publication No. US20020061515A1
; GENERAL INFORMATION:
; APPLICANT: Lynch, Gary
; APPLICANT: Bi, Xiaoning
; APPLICANT: Gall, Christine M.
; TITLE OF INVENTION: Model for Neurodegenerative Diseases Involving Amyloid
; TITLE OF INVENTION: Accumulation
; FILE REFERENCE: 1819.0040001
; CURRENT APPLICATION NUMBER: US/09/961,381A
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,374
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-961-381A-3

Query Match 100.0%; Score 34; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6
DB 1 GRGDSP 6

RESULT 13
US-09-380-682-63
; Sequence 63, Application US/09380682
; Publication No. US20030064431A1
; GENERAL INFORMATION:
; APPLICANT: MOSSAKOWSKA, Danuta Ewa Irena
; APPLICANT: COX, Vivienne Frances
; APPLICANT: SMITH, Richard Anthony Godwin
; TITLE OF INVENTION: COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE
; SEQUENCES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington

```

STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.33  
CURRENT APPLICATION DATA: US/C9/380,682  
APPLICATION NUMBER: US/C9/380,682  
FILING DATE: 19-Oct-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB98/00727  
FILING DATE: 05-MAR-1998  
APPLICATION NUMBER: GB 9704519.9  
FILING DATE: 05-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: ISACSON, John P.  
REGISTRATION NUMBER: 33,715  
REFERENCE/DOCKET NUMBER: 88362/137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-09-380-682-63

Query Match 100.0%; Score 34; DB 12; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 14  
US-10-136-187-42  
; Sequence 42, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: C'more, Sally  
; APPLICANT: C'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-136-187-42

Query Match 100.0%; Score 34; DB 12; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

RESULT 15  
US-10-001-945-31  
; Sequence 31, Application US/10001945  
; Publication No. US20020151493A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, Gary L.  
; APPLICANT: Self, Christopher  
; APPLICANT: Lee, Lily  
; APPLICANT: Cook, Charles M.  
; APPLICANT: Birktope, Jens  
; TITLE OF INVENTION: THERAPEUTIC AGENTS AND METHODS OF USE THEREOF FOR THE  
; FILE REFERENCE: PPI-106CP2  
; CURRENT APPLICATION NUMBER: US/10/001,945  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 09/972,772  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: US 09/704,251  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Motifs  
US-10-001-945-31

Query Match 100.0%; Score 34; DB 13; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
Db 1 GRGDSP 6

Search completed: April 16, 2004, 08:03:16  
Job time : 40.35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 12.6 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-7

Perfect score: 34

Sequence: 1 GRGDSP 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	273	2 A28512	fibronectin - chic
2	34	100.0	441	2 E84264	isochorismate synt
3	34	100.0	691	2 T45933	hypothetical prote
4	34	100.0	1020	2 A29355	fibronectin - chic
5	34	100.0	2265	1 FNBO	fibronectin - bovi
6	34	100.0	2386	1 FNHD	fibronectin precu
7	34	100.0	2477	2 S14428	fibronectin precu
8	34	100.0	2481	2 A43908	fibronectin - Afri
9	31	91.2	352	2 A60965	vinculin - mouse
10	31	91.2	370	2 S27344	mapK protein - rhi
11	31	91.2	450	2 D87342	multidrug resist
12	31	91.2	1066	2 T10108	vinculin - mouse
13	31	91.2	1134	1 A35955	meta-vinculin - hu
14	31	91.2	1135	1 A29997	meta-vinculin - ch
15	31	91.2	1307	2 T30887	146D nuclear prote
16	30	88.2	95	2 T45263	cobalt transport p
17	30	88.2	232	2 A87363	dnaf family protei
18	30	88.2	235	1 B41326	nitrile hydratase
19	30	88.2	294	2 H83544	probable transcrip
20	30	88.2	301	2 T358C5	hypothetical prote
21	30	88.2	511	2 JF03112	3-ketosteroid-deh
22	30	88.2	522	2 T37320	probable dehydroge
23	30	88.2	625	2 A34615	profilaggrin - rat
24	30	88.2	633	2 T31160	maturase-related p
25	30	88.2	682	2 T28899	hypothetical prote
26	30	88.2	710	2 T49516	Atu related protei
27	30	88.2	714	2 A83595	ATP-dependent DNA
28	30	88.2	715	2 B70741	probable moey prot
29	30	88.2	825	2 JC4163	DNA-binding protei

30	88.2	846	2 A12336	primosomal protein
31	88.2	950	2 T44422	oxoglutarate dehyd
32	88.2	969	2 A75634	McrB-related prote
33	88.2	1188	2 E96642	hypothetical prote
34	88.2	1265	1 A37967	neural cell adhesi
35	88.2	1272	2 S26180	neurofascin - chic
36	88.2	1453	2 F88640	protein F52C12.4 (
37	88.2	1453	2 S17384	T-cell receptor be
38	85.3	72	2 A42856	EPF autoantibody-1
39	85.3	267	2 F90516	esterase/lipase 1
40	85.3	294	2 A05071	pol polyprotein -
41	85.3	387	2 F87268	recF protein (impo
42	85.3	424	2 S71837	vasopressin recept
43	85.3	428	2 E87463	hypothetical prote
44	85.3	490	2 A29782	unspecific monooxy
45	85.3	514	2 H83868	hypothetical prote

## ALIGNMENTS

### RESULT 1

A28512  
fibronectin - chicken (fragment)  
C:Species: Gallus gallus (Chicken)  
C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Aug-1999  
C:Accession: A28512  
R:Kubomura, S.; Obara, M.; Karasaki, Y.; Taniguchi, H.; Gotoh, S.; Tsuda, T.; Higashi, K.  
Biochim. Biophys. Acta 910, 171-181, 1987  
A>Title: Genetic analysis of the cell binding domain region of the chicken fibronectin g  
A:Reference number: A28512; MUID:88050950; PMID:2823899  
A:Accession: A28512  
A:Molecule type: DNA  
A:Residues: 1-273 <KUB>  
A:Cross-references: GB:X06533; NID:963393; PIDN:CAA29781.1; PID:9295716  
A>Note: the authors translated the codon CCG for residue 190 as Gln, CAG for residue 24:  
C:Genetics:  
A:Introns: 90/1; 129/1; 184/1; 236/1  
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe  
C:Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; heter  
F:1-82/Domain: fibronectin type III repeat homology (fragment) <FN3I>  
F:90-172/Domain: fibronectin type III repeat homology <FN3J>  
F:167-169/Region: cell attachment (R-G-D) motif  
F:184-268/Domain: fibronectin type III repeat homology <FN3K>

Query Match 100.0%; Score 34; DB 2; Length 273;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 5

DB 166 GRGDSP 171

### RESULT 2

E84264  
isochorismate synthase [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 12-Jun-2003  
C:Accession: E84264  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
Jung, K.H.; Alam, W.; Freitas, F.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A>Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: E84264  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-441 <STC>  
A:Cross-references: GB:AE004437; NID:gi0580629; PIDN:AA019481.1; GSPDB:GN00138  
C:Genetics:

A:Gene: menF  
C:Superfamily: isochorismate synthase

Query Match 100.0%; Score 34; DB 2; Length 441;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6  
DB 271 GRGDSF 276

RESULT 3  
T45933  
hypothetical protein F5K20.230 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 34-Feb-2000  
C:Accession: T45933  
R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.  
A:Reference number: Z23C17  
A:Accession: T45933  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-431  
A:Cross-references: EMBL:AL132960  
A:Experimental source: Cultivar Columbia; BAC clone F5K20  
C:Genetics:  
A:Map position: 3  
A:Introns: 87/3; 127/1; 146/3; 175/2; 206/3; 243/1; 267/2; 391/1; 511/3; 594/1; 617/3  
A:Note: F5K20.230

Query Match 100.0%; Score 34; DB 2; Length 691;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6  
DB 389 GRGDSF 394

RESULT 4  
A29355  
fibronectin - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 12-Feb-1999  
C:Accession: A29355  
R:Korton, F.A.; Hynes, R.O.  
Mol. Cell. Biol. 7, 4297-4307, 1987  
A:Title: Alternative splicing of chicken fibronectin in embryos and in normal and trans-  
A:Reference number: A29355; MUID:88142820; PMID:2830487  
A:Accession: A29355  
A:Molecule type: mRNA  
A:Residues: 1-1020 <NOR>  
C:Genetics:  
A:Introns: 176/3  
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat  
C:Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; heter  
F:1-86/Domain: fibronectin type III repeat homology <FN3H>  
F:92-177/Domain: fibronectin type III repeat homology <FN3J>  
F:180-262/Domain: fibronectin type III repeat homology <FN3L>  
F:257-259/Region: cell attachment (R-G-D) motif  
F:274-356/Domain: fibronectin type III repeat homology <FN3K>  
F:364-446/Domain: fibronectin type III repeat homology <FN3M>  
F:454-536/Domain: fibronectin type III repeat homology <FN3N>  
F:546-628/Domain: fibronectin type III repeat homology <FN3O>  
F:636-718/Domain: fibronectin type III repeat homology <FN3P>  
F:837-917/Domain: fibronectin type III repeat homology <FN3Q>  
F:940-979/Domain: fibronectin type III repeat homology <FN3R>  
F:940-963, 967-979, 985-1012/Disulfide bonds: #status predicted

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSF 6  
DB 256 GRGDSF 261

RESULT 5  
FNBO  
fibronectin - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Oct-2000  
C:Accession: A26452; B21165; A23292  
R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.  
Eur. J. Biochem. 161, 441-453, 1986  
A:Title: Complete primary structure of bovine plasma fibronectin.  
A:Reference number: A26452; MUID:87054047; PMID:3780752  
A:Accession: A26452  
A:Molecule type: protein  
A:Residues: 1-2265 <SKO>  
R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983  
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin  
A:Reference number: A21165; MUID:83221567; PMID:6304699  
A:Accession: B21165  
A:Molecule type: mRNA  
A:Residues: 2170-2265 <KOR>  
A:Cross-references: GB:K00800; NID:6163055; PID:AAA30521.2; PID:G5713323  
R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sott  
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983  
A:Title: Partial primary structure of bovine plasma fibronectin: three types of interna  
A:Reference number: A23292; MUID:83117805; PMID:6218503  
A:Accession: A23292  
A:Molecule type: protein  
A:Residues: 1-16, 'C', 18-20, 'S', 22-432, 447-463, 1367-1517, 1567-1673, 2062-2176, 'N', 2178-22  
C:Comment: Cys-1201 and Cys-2015 have free sulphydryl groups.  
C:Comment: The plasma fibronectin molecule consists of two chains, which are connected  
C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fi  
aling, and maintenance of cell shape.  
C:Comment: Plasma fibronectin is synthesized by hepatocytes.  
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep  
C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracell  
F:21-56/Domain: fibronectin type I repeat homology <IF1>  
F:21-241/Domain: fibronectin type I repeat homology <IF2>  
F:66-104/Domain: fibronectin type I repeat homology <IF3>  
F:110-148/Domain: fibronectin type I repeat homology <IF4>  
F:155-194/Domain: fibronectin type I repeat homology <IF5>  
F:200-239/Domain: fibronectin type I repeat homology <IF6>  
F:277-577/Domain: collagen binding <CBR>  
F:277-311/Domain: fibronectin type I repeat homology <IF7>  
F:329-370/Domain: fibronectin type II repeat homology <2F1>  
F:389-430/Domain: fibronectin type II repeat homology <2F2>  
F:439-527/Domain: fibronectin type I repeat homology <IF8>  
F:487-524/Domain: fibronectin type I repeat homology <IF9>  
F:530-568/Domain: fibronectin type III repeat homology <FN3A>  
F:578-681/Domain: fibronectin type III repeat homology <FN3B>  
F:688-770/Domain: fibronectin type III repeat homology <FN3C>  
F:779-860/Domain: fibronectin type III repeat homology <FN3D>  
F:875-957/Domain: fibronectin type III repeat homology <FN3E>  
F:965-1046/Domain: fibronectin type III repeat homology <FN3F>  
F:1055-1134/Domain: fibronectin type III repeat homology <FN3G>  
F:1142-1227/Domain: fibronectin type III repeat homology <FN3H>  
F:1235-1318/Domain: fibronectin type III repeat homology <FN3I>  
F:1326-1404/Domain: fibronectin type III repeat homology <FN3J>  
F:1410-1517/Domain: cell attachment <CAD>  
F:1416-1502/Domain: fibronectin type III repeat homology <FN3L>  
F:1493-1495/Region: cell attachment (R-G-D) motif  
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>  
F:1600-1870/Domain: heparin binding <HB2>  
F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>  
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>  
F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>  
F:1970-1972/Region: cell attachment (R-G-D) motif

F:1982-2062/Domain: fibronectin type III repeat homology <FN30>  
F:1985-2216/Domain: fibrin binding <FB2>  
F:2085-2124/Domain: fibronectin type I repeat homology <IF1C>  
F:2130-2167/Domain: fibronectin type I repeat homology <IF1I>  
F:2174-2209/Domain: fibronectin type I repeat homology <IF12>  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of fibrin) #status experimental  
F:21-47,45-56,66-94,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-3  
7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted  
F:339,497,511,846,976,1213,1987/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:1943,1944/Binding site: carboxylate (Thr) (covalent) #status experimental  
F:2246/Disulfide bonds: interchain (to 2252) #status predicted  
F:2250/Disulfide bonds: interchain (to 2246) #status predicted  
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 34; DB 1; Length 2265;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GRGDSP 6  
|||||  
Db 1492 GRGDSP 1497

RESULT 6  
FNHU

N:Alternate names: fibronectin splice form ED-A  
C:Species: Homo sapiens (man)  
C:Date: 27-Nov-1985 #sequence, revision 31-Mar-1993 #text change 08-Dec-2000  
C:Accession: A26460; A26284; S03917; A24854; A24476; A93529; A21011; A90495; A22  
R:Dean, D.C.; Bowlus, C.E.; Bourgeois, S.  
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987  
A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.  
A:Reference number: A26460; MUID:87175578; PMID:3031656  
A:Accession: A26460  
A:Molecule type: DNA  
A:Residues: 1-49 <DEA>  
A:Cross-references: GB:M15801; NID:G182686; PIDN:AAA53376.1; PID:G553293  
R:Oldberg, A.; Ruoslahti, E.  
J. Biol. Chem. 261, 2113-2116, 1986  
A:Title: Evolution of the fibronectin gene.  
A:Reference number: A26284; MUID:86111901; PMID:3003095  
A:Accession: A26284  
A:Molecule type: DNA  
A:Residues: 1447-1540 <OLD>  
A:Cross-references: GB:M12549; NID:G182688  
A:Note: the authors translated the codon TTC for residue 1494 as Glu  
R:Paolella, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.  
Nucleic Acids Res. 16, 3545-3557, 1988  
A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B  
A:Reference number: S00848; MUID:88233940; PMID:3375063  
A:Accession: S03917  
A:Molecule type: DNA  
A:Residues: 1594-1767, 'V', 1769-1783 <PAO>  
A:Cross-references: EMBL:X07718; NID:G31402  
A:Note: the authors translated the codon AAC for residue 1631 as Asp  
R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.  
FEBS Lett. 207, 287-291, 1986  
A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.  
A:Reference number: A24854; MUID:87030929; PMID:3770201  
A:Accession: A24854  
A:Molecule type: DNA  
A:Residues: 1932-2147 <VIB>  
A:Cross-references: GB:X34530; NID:G31436  
R:Gutman, A.; Yamada, K.M.; Kornblitt, A.  
FEBS Lett. 207, 145-148, 1986  
A:Title: Human fibronectin is synthesized as a pre-propolypeptide.  
A:Reference number: A24476; MUID:87030890; PMID:3770189  
A:Accession: A24476  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA

A:Residues: 1-14, 'Q', 16-38 <GUT>  
R:Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.  
EMBO J. 4, 1755-1759, 1985  
A:Title: Primary structure of human fibronectin: differential splicing may generate at l  
A:Reference number: A91008; MUID:85284965; PMID:2992939  
A:Accession: A91008  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 32-1344, 1346-2080, 2112-2386 <K03>  
A:Cross-references: GB:X02761  
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
Nucleic Acids Res. 12, 5853-5868, 1984  
A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptic  
A:Reference number: A93529; MUID:84272258; PMID:6462919  
A:Accession: A93529  
A:Molecule type: mRNA  
A:Residues: 973-2080; 2112-2386 <K02>  
A:Cross-references: GB:X00739  
R:Oldberg, A.; Linney, E.; Ruoslahti, E.  
J. Biol. Chem. 258, 10193-10196, 1983  
A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell s  
A:Reference number: A21011; MUID:83290929; PMID:6688418  
A:Accession: A21011  
A:Molecule type: mRNA  
A:Residues: 1434-1537 <OL2>  
A:Cross-references: GB:K00055; NID:G182680; PIDN:AAA52459.1; PID:G182683  
R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.  
Biochemistry 24, 2698-2704, 1985  
A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra  
A:Reference number: A90495; MUID:85280409; PMID:2992573  
A:Accession: A90495  
A:Molecule type: mRNA  
A:Residues: 1534-2396 <BER>  
A:Cross-references: GB:M10905; NID:G182696; PIDN:AAA52462.1; PID:G182697  
R:Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.  
FEBS Lett. 186, 31-34, 1985  
A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.  
A:Reference number: A22245; MUID:85231203; PMID:2989004  
A:Accession: A22245  
A:Molecule type: mRNA  
A:Residues: 1948-2067 <UME>  
A:Cross-references: GB:M27589; NID:G182705; PIDN:AAA52465.1; PID:G182706  
A:Accession: B22245  
A:Molecule type: mRNA  
A:Residues: 1975-1991, 2017-2039 <UM2>  
A:Cross-references: GB:M27590; NID:G182701; PIDN:AAA52464.1; PID:G182704  
R:Seiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.  
Biochemistry 25, 4936-4941, 1986  
A:Title: Human liver fibronectin complementary DNAs: identification of two different mes  
A:Reference number: 152394; MUID:87026578; PMID:3021206  
A:Accession: 165273  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>  
A:Cross-references: GB:M14060; NID:G182701; PIDN:AAA52464.1; PID:G182704  
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983  
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins  
A:Reference number: A21165; MUID:83221567; PMID:6304599  
A:Accession: A21165  
A:Molecule type: mRNA  
A:Residues: 2291-2386 <K03>  
A:Cross-references: GB:X00799; NID:G182681; PIDN:AAA52460.1; PID:G182684  
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.  
J. Biol. Chem. 258, 12670-12674, 1983  
A:Title: Primary structure of human plasma fibronectin.  
A:Reference number: A92398; MUID:84032463; PMID:6630202  
A:Accession: A92398  
A:Molecule type: protein  
A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>  
R:Garcia-Pardo, A.; Gold, L.I.  
Arch. Biochem. Biophys. 304, 181-188, 1993  
A:Title: Further characterization of the binding of fibronectin to gelatin reveals the p

A:Reference number: S34791; MUID:93312001; PMID:8323285  
A:Accession: S34791  
A:Molecule type: protein  
A:Residues: 291-300:552-560 <GAR2>  
R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.  
Thromb. Res. 43, 469-477, 1986  
A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.  
A:Reference number: A60904; MUID:87019725; PMID:3532418  
A:Accession: A60904  
A:Molecule type: protein  
A:Residues: 293-301 <GRI>  
R:Calaycay, J.; Pande, H.; Lee, T.; Borsl, L.; Siri, A.; Shively, J.E.; Zardi, L.  
J. Biol. Chem. 260, 12136-12141, 1985  
A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl  
A:Reference number: A23901; MUID:86008277; PMID:3900070  
A:Accession: A23901  
A:Molecule type: protein  
A:Residues: 616-677, 'Q', 679-703, 'PT' <CAL>  
R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.  
J. Biol. Chem. 257, 9593-9597, 1982  
A:Title: The cell attachment domain of fibronectin. Determination of the primary structu  
A:Reference number: A92386; MUID:82265604; PMID:7050098  
A:Accession: A92386  
A:Molecule type: protein  
A:Residues: 1441-1548 <PIE>  
A:Note: residues 1524-1527 are responsible for the cell-binding activity  
R:Garcia-Pardo, A.; Rostagno, A.; Frangione, B.  
Biochem. J. 241, 923-928, 1987  
A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom  
A:Reference number: A32517; MUID:87241275; PMID:3593230  
A:Accession: A32517  
A:Molecule type: protein  
A:Residues: 1589-1630, 'T', 1722-2058 <GAR3>  
R:Ressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand  
Biochem. J. 274, 731-738, 1991  
A:Title: Human plasma fibronectin. Demonstration of structural differences between the A  
A:Reference number: S14357; MUID:91190085; PMID:2012601  
A:Accession: S14357  
A:Molecule type: protein  
A:Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>  
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.  
J. Biol. Chem. 260, 10320-10325, 1985  
A:Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal  
A:Reference number: A23891; MUID:85261459; PMID:4019516  
A:Accession: A23891  
A:Molecule type: protein  
A:Residues: 2071-2080; 2112-2356 <GAR4>  
C:Comment: The extra domain and connecting strand 3 are subject to developmental and tis  
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,  
ation, and transformation.  
C:Genetics:  
A:Gene: GDS:FN1  
A:Cross-references: GDB:119135; OMIM:135600  
A:Map position: 2q34-q34  
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep  
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat  
F1-26/Domain: signal sequence #status predicted <SIG>  
F127-31/Domain: propeptide #status predicted <PRO>  
F132-2386/Product: fibronectin #status experimental <MAT>  
F152-272/Domain: fibrin and heparin binding <PHB>  
F152-87/Domain: fibronectin type I repeat homology <IF1>  
F197-135/Domain: fibronectin type I repeat homology <IF2>  
F141-179/Domain: fibronectin type I repeat homology <IF3>  
F186-225/Domain: fibronectin type I repeat homology <IF4>  
F231-270/Domain: fibronectin type I repeat homology <IF5>  
F308-608/Domain: collagen binding <CBR>  
F308-342/Domain: fibronectin type I repeat homology <IF6>  
F360-401/Domain: fibronectin type II repeat homology <2F2>  
F420-461/Domain: fibronectin type II repeat homology <2F3>  
F470-508/Domain: fibronectin type I repeat homology <IF7>  
F518-555/Domain: fibronectin type I repeat homology <IF8>  
F561-599/Domain: fibronectin type I repeat homology <IF9>

F:609-692/Domain: fibronectin type III repeat homology <3FA>  
F:616-706/Domain: heparin binding <HEB>  
F:719-801/Domain: fibronectin type III repeat homology <3FB>  
F:810-891/Domain: fibronectin type III repeat homology <3FC>  
F:906-988/Domain: fibronectin type III repeat homology <3FD>  
F:996-1077/Domain: fibronectin type III repeat homology <3FE>  
F:1086-1164/Domain: fibronectin type III repeat homology <3FF>  
F:1173-1258/Domain: fibronectin type III repeat homology <3FG>  
F:1266-1349/Domain: fibronectin type III repeat homology <3FH>

Query Match 100.0%; Score 34; DB 1; Length 2386;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
| | | | |  
DB 1523 GRGDSP 1528

RESULT 7  
S14428  
fibronectin precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 20-Aug-1999  
C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049  
R:Hynes, R.O.  
Submitted to the EMBL Data Library, July 1989  
A:Reference number: S14428  
A:Accession: S14428  
A:Molecule type: mRNA  
A:Residues: 1-2477 <HYN>  
A:Cross-references: EMBL:X15906; MID:G56163; PID:CAA34020.1; PID:G56164  
R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.  
EMBO J. 6, 2573-2580, 1987  
A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.  
A:Reference number: S12455; MUID:88054951; PMID:2445560  
A:Accession: S12455  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 609-1810, 'T', 1812-2283 <SCH>  
A:Cross-references: EMBL:X15906  
R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984  
A:Title: A single rat fibronectin gene generates three different mRNAs by alternative sp  
A:Reference number: A22319; MUID:84298097; PMID:6089177  
A:Accession: A22319  
A:Molecule type: DNA  
A:Residues: 2052-2237 <TAM>  
R:Falkenberg, C.; Englund, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstrøm, B.  
Biochem. J. 301, 745-751, 1994  
A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex i  
A:Reference number: S46203; MUID:94330948; PMID:7519849  
A:Accession: S46203  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1183-1192 <GLN>, 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <FAL>  
R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.  
EMBO J. 6, 2565-2572, 1987  
A:Title: Organization of the fibronectin gene provides evidence for exon shuffling durin  
A:Reference number: S00459; MUID:88054950; PMID:3119323  
A:Accession: S00459  
A:Molecule type: DNA  
A:Residues: 1-139; 2382-2477 <PAT>  
A:Cross-references: EMBL:X05831  
A:Note: The authors translated the codon CCT for residues 51 and 94 as Ala  
R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.  
Cell 35, 421-431, 1983  
A:Title: Three different fibronectin mRNAs arise by alternative splicing within the cod.  
A:Reference number: A27252; MUID:84082067; PMID:6317187  
A:Accession: A27252  
A:Molecule type: mRNA  
A:Residues: 1586-1720, 'T', 1722, 1813-2477 <SC2>  
R:Odermatt, E.; Tamkun, J.W.; Hynes, R.O.



Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6

Db 337 GKGDSF 342

RESULT 10

S27344

hupK protein - Rhizobium leguminosarum

C:Species: Rhizobium leguminosarum

C>Date: 25-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 22-Jun-2003

C:Accession: S27344

R:Rey, L.; Hidalgo, E.; Palacios, J.; Ruiz-Argueso, T.

J. Mol. Biol. 228, 998-1002, 1992

A>Title: Nucleotide sequence and organization of an H(2)-uptake gene cluster from Rhizob

A:Reference number: S27340; MUID:93108466; PMID:1469733

A:Accession: S27344

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-370 <REV>

A:Cross-references: EMBL:X52974; NID:gl167855; PIDN:CAA37158.1; PID:948731

C:Superfamily: hydrogenase maturation factor, hupK type

Query Match 91.2%; Score 31; DB 2; Length 370;

Best Local Similarity 83.3%; Pred. No. 53;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6

Db 163 GRGDTP 168

RESULT 11

D87342

multidrug resistance protein, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: D87342

R:Niermar, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: D87342

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-450 <STO>

A:Cross-references: GB:AE005673; NID:gl3421986; PIDN:AAK22736.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0751

Query Match 91.2%; Score 31; DB 2; Length 450;

Best Local Similarity 83.3%; Pred. No. 65;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6

Db 220 GRGDTP 225

RESULT 12

T10108

vinculin - mouse

C:Species: Mus musculus (house mouse)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: T10108

R:Alatorov, V.E.; Kramerova, I.A.; Frolov, M.V.; Lavrov, S.A.; Westphal, B.D.

FEBS Lett. 413, 197-201, 1997

A>Title: Vinculin gene is non-essential in Drosophila melanogaster.

A:Reference number: Z16953; MUID:97424359; PMID:9280281

A:Accession: T10108

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1066 <HLA>

A:Cross-references: EMBL:L18880; NID:g309532; PIDN:AAB96843.1; PID:g309533

A:Experimental source: strain 129SVJ

C:Superfamily: meta-vinculin; vinculin amino-terminal homology; vinculin carboxyl-termi

C:Keywords: cell adhesion; cytoskeleton

F:4-257/Domain: vinculin amino-terminal homology <VNN>

Query Match 91.2%; Score 31; DB 2; Length 1066;

Best Local Similarity 83.3%; Pred. No. 1.7e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6

Db 452 GRGDSP 457

RESULT 13

A35955

meta-vinculin - human

N:Contains: vinculin

C:Species: Homo sapiens (man)

C>Date: 09-Nov-1990 #sequence\_revision 02-Aug-1996 #text\_change 22-Jun-1999

C:Accession: A35955; S39760; S21180; S39758; S24424

R:Weller, P.A.; Ogryzko, E.P.; Corben, E.B.; Zhidkova, N.I.; Patel, B.; Price, G.J.; Sp

Proc. Natl. Acad. Sci. U.S.A. 87, 5667-5671, 1990

A>Title: Complete sequence of human vinculin and assignment of the gene to chromosome 11

A:Reference number: A35955; MUID:90332642; PMID:2116004

A:Accession: A35955

A:Molecule type: mRNA

A:Residues: 1-915,984-1134 <WEL>

A:Cross-references: GB:M33308; NID:g340236; PIDN:AAA61283.1; PID:g340237

A>Note: the authors translated the codon GAG for residue 60 as Glu

R:Korelansky, V.E.; Ogryzko, E.P.; Zhidkova, N.I.; Weller, P.A.; Critchley, D.R.; Vanc

Eur. J. Biochem. 204, 767-772, 1992

A>Title: An additional exon in the human vinculin gene specifically encodes meta-vincul

lin insert.

A:Reference number: S21180; MUID:92174935; PMID:1339348

A:Accession: S39760

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 854-1051 <KOT>

A:Cross-references: GB:S87180; NID:g246706; PIDN:AAB21656.1; PID:g246707

A:Accession: S21180

A:Molecule type: mRNA

A:Residues: 901-995 <KO3>

A:Accession: S39758

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 911-983 <KO2>

A>Note: meta-vinculin isoform is specific to muscle

C:Genetics:

A:Gene: GDB:VCL

A:Cross-references: GDB:125348; OMIM:193065

A:Map position: 10q22.1-10q23

A:Introns: 915/3; 983/3

C:Superfamily: meta-vinculin; vinculin amino-terminal homology; vinculin carboxyl-termi

C:Keywords: alternative splicing; cell adhesion; cytoskeleton; duplication; lipid bindi

F:1-1134/Product: meta-vinculin #status predicted <MAL>

F:1-915,984-1134/Product: vinculin #status predicted <MAT>

F:4-257/Domain: vinculin amino-terminal homology <VINN>

F:167-207/Region: talin binding

F:259-589/Region: 112-residue repeats

F:581-1132/Domain: vinculin carboxyl-terminal homology <VINC>

F:837-878/Region: proline-rich

F:822/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 1; Length 1134;

Best Local Similarity 83.3%; Pred. No. 1.8e-02;

Matches 5; Conservative 1; Mismatches 0; Indels 3; Gaps 0;

QY 1 GRGDSP 6



F;837-878/Region: proline-rich  
F;822/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 91.2% Score 31; DB 1; Length 1135;  
Best Local Similarity 83.3% Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
Db 452 GKGDSP 457

## RESULT 15

T30887  
146D nuclear protein - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
C;Accession: T30887  
R;Schmidt-Zachmann, M.S.; Knecht, S.; Kraemer, A.  
Mol. Biol. Cell 9, 143-160, 1998  
A;Title: Molecular characterization of a novel, widespread nuclear protein that co-local  
A;Reference number: Z20921; MUID:98099695; PMID:9436997  
A;Accession: T30887  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1307 <SCH>  
A;Cross-references: EMBL:Y08997; NID:g2791895; PIDN:CAA70201.1; PID:g2791896

Query Match 91.2% Score 31; DB 2; Length 1307;  
Best Local Similarity 83.3% Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
Db 272 GKGDSP 277

Search completed: April 16, 2004, 07:29:22  
Job time : 13.6 secs

Db 452 GKGDSP 457  
|:||||

## RESULT 14

A29997  
meta-vinculin - chicken  
N;Contains: vinculin  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1988 #sequence\_revision 02-Aug-1996 #text\_change 22-Jun-1999  
C;Accession: A31346; S03973; A27884; S39759; S32807; S65691; A29997  
R;Coutu, M.D.; Craig, S.W.  
Proc. Natl. Acad. Sci. U.S.A. 35, 8535-8539, 1988  
A;Title: cDNA-derived sequence of chicken embryo vinculin.  
A;Reference number: A31346; MUID:890422-6; PMID:3141928  
A;Accession: A31346  
A;Molecule type: mRNA  
A;Residues: 1-915, 985-1135 <COU>  
A;Cross-references: GB:J04126; NID:g212872; PIDN:AAA49136.1; PID:g212873  
R;Price, G.J.; Jones, P.; Davison, M.D.; Patel, B.; Bendori, R.; Geiger, B.; Critchley,  
Biochem. J. 259, 453-461, 1989  
A;Title: Primary sequence and domain structure of chicken vinculin.  
A;Reference number: S03973; MUID:89246350; PMID:2497736  
A;Accession: S03973  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-441, 'QLSQCI', 448-700, 'H', 702-915, 985-1135 <PRI>  
A;Note: this sequence has been revised in reference A35955  
R;Price, G.J.; Jones, P.; Davison, M.D.; Patel, B.; Eperon, I.C.; Critchley, D.R.  
Biochem. J. 245, 595-603, 1987  
A;Title: Isolation and characterization of a vinculin cDNA from chick-embryo fibroblasts  
A;Reference number: A27884; MUID:88024106; PMID:3117046  
A;Accession: A27884  
A;Molecule type: mRNA  
A;Residues: 1-441, 'QLSQCI', 448-700, 'H', 702-879, 'K', 881 <PR2>  
A;Note: this sequence has been revised in reference A35955  
R;Weller, P.A.; Ogryzko, E.P.; Corben, E.B.; Zhidkova, N.I.; Patel, B.; Price, G.J.; Spu  
Proc. Natl. Acad. Sci. U.S.A. 87, 5667-5671, 1990  
A;Title: Complete sequence of human vinculin and assignment of the gene to chromosome 1q  
A;Reference number: A35955; MUID:90332642; PMID:2116004  
R;Kotliansky, V.E.; Ogryzko, E.P.; Zhidkova, N.I.; Weller, P.A.; Critchley, D.R.; Vance  
Eur. J. Biochem. 204, 767-772, 1992  
A;Title: An additional exon in the human vinculin gene specifically encodes meta-vinculin  
lin insert.  
A;Reference number: S21180; MUID:92174935; PMID:1339348  
A;Accession: S39759  
A;Molecule type: mRNA  
A;Residues: 915-984 <KCT>  
R;Byrne, B.J.; Kaczorowski, Y.J.; Coutu, M.D.; Craig, S.W.  
J. Biol. Chem. 267, 12845-12850, 1992  
A;Title: Chicken vinculin and meta-vinculin are derived from a single gene by alternativ  
A;Reference number: S32807; MUID:92316975; PMID:1618784  
A;Accession: S32807  
A;Molecule type: DNA  
A;Residues: 916-943, 'S', 945-984 <BYR>  
A;Cross-references: EMBL:M87837; NID:g212870; PIDN:AAA49135.1; PID:g212871  
A;Accession: S65691  
A;Molecule type: mRNA  
A;Residues: 916-943, 'S', 945-984 <BYX>  
A;Cross-references: EMBL:M87837; NID:g212870; PIDN:AAA49135.1; PID:g212871  
A;Note: meta-vinculin isoform is specific to muscle  
C;Genetics:  
A;Gene: vinc1  
A;Map position: 14  
C;Superfamily: meta-vinculin; vinculin amino-terminal homology; vinculin carboxyl-termin  
C;Keywords: alternative splicing; cell adhesion; cytoskeleton; duplication; lipid bindin  
F;1-1135/Product: meta-vinculin #status predicted <MA>  
F;1-915, 985-1135/Product: vinculin #status predicted <MA2>  
F;4-257/Domain: vinculin amino-terminal homology <VINN>  
F;167-207/Region: talin binding  
F;259-589/Region: 112-residue repeats  
F;581-1133/Domain: vinculin carboxyl-terminal homology <VINC>

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time 10.5 Seconds  
(without alignments)  
29.754 Million cell updates/sec

Title: US-09-991-588B-7

Perfect score: 34

Sequence: 1 GRGDSF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	1256	1	F11722 gallus gall
2	34	100.0	1328	1	Q91289 pleurodeles
3	34	100.0	2265	1	F1NC_BOVIN
4	34	100.0	2386	1	P02751 homo sapien
5	34	100.0	2477	1	F1NC_MOUSE
6	34	100.0	2477	1	F1NC_MOUSE
7	34	100.0	2481	1	F1NC_XENLA
8	31	91.2	370	1	HUPK_RHILV
9	31	91.2	1030	1	SMGA_HUMAN
10	31	91.2	1065	1	V1NC_CHICK
11	31	91.2	1065	1	V1NC_MOUSE
12	31	91.2	1133	1	V1NC_HUMAN
13	31	91.2	1304	1	S3B1_HUMAN
14	31	91.2	1304	1	S3B1_MOUSE
15	31	91.2	1307	1	S3B1_XENLA
16	30	88.2	95	1	CBIN_METTM
17	30	88.2	235	1	NHAB_RHOSC
18	30	88.2	343	1	ADP4_STRCO
19	30	88.2	715	1	YD55_MYTCT
20	30	88.2	825	1	5E5_RAT
21	30	88.2	950	1	OD01_ALCEU
22	30	88.2	1266	1	NGCA_CHICK
23	30	88.2	1369	1	NFAS_CHICK
24	29	85.3	247	1	STC1_MOUSE
25	29	85.3	247	1	STC1_RAT
26	29	85.3	294	1	POL_SMSAV
27	29	85.3	387	1	REC7_CAUCR
28	29	85.3	424	1	V1AR_RAT
29	29	85.3	490	1	CPG8_HUMAN
30	29	85.3	503	1	AMPA_COREF
31	29	85.3	582	1	PUR6_CRYNE
32	29	85.3	585	1	DHSA_BACSU
33	29	85.3	591	1	NOS2_CANFA

34	29	85.3	601	1	TOPM_HUMAN
35	29	85.3	829	1	TOP1_XENLA
36	29	85.3	883	1	HSS2_HUMAN
37	29	85.3	883	1	HSS2_MOUSE
38	29	85.3	1136	1	NOS2_CHICK
39	29	85.3	1242	1	JAG1_BRARE
40	28	82.4	85	1	Y522_PSEAE
41	28	82.4	210	1	KCY_AGRTS
42	28	82.4	212	1	KCY_RHIME
43	28	82.4	215	1	HIS5_DEIRA
44	28	82.4	216	1	KCY_RHILO
45	28	82.4	232	1	FLGH_CANJE

## ALIGNMENTS

RESULT 1  
F1NC\_CHICK STANDARD; PRT; 1256 AA.  
AC P11722; Q90921;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Fibronectin (FN) (Fragments).  
GN FNI.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP MEDLINE=83117850; PubMed=6572007;  
RA Hirano H., Yamada Y., Sullivan M., de Crombrughe B., Pastan I.,  
RA Yamada K.M.;  
RT "Isolation of genomic DNA clones spanning the entire fibronectin  
gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:46-50(1983).  
RN [2]  
RP SEQUENCE OF 51-1256 FROM N.A.  
RC STRAIN=White leghorn;  
RA Norton P.A.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 227-415 FROM N.A.  
RX MEDLINE=96183658; PubMed=8603103;  
RA Gehris A.L., Brandli D.W., Lewis S.D., Bennett V.D.;  
RT "The exon encoding the fibronectin type III-9 repeat is  
constitutively included in the mRNA from chick limb mesenchyme and  
cartilage.";  
RL Biochim. Biophys. Acta 1311:5-12(1996).  
RN [4]  
RP SEQUENCE OF 327-599 FROM N.A.  
RX MEDLINE=88050950; PubMed=2823899;  
RA Kubomura S., Obara M., Karasaki Y., Taniguchi H., Gotoh S.,  
RA Tsuda T., Higashi K., Ohsato K., Hiarno H.;  
RT "Genetic analysis of the cell binding domain region of the chicken  
fibronectin gene.";  
RL Biochim. Biophys. Acta 910:171-181(1987).  
RN [5]  
RP SEQUENCE OF 413-1256 FROM N.A.  
RX MEDLINE=88142820; PubMed=2830487;  
RA Norton P.A., Hynes R.O.;  
RT "Alternative splicing of chicken fibronectin in embryos and in normal  
and transformed cells.";  
RL Mol. Cell. Biol. 7:4297-4307(1987).  
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds  
including collagen, fibrin, heparin, DNA, and actin. Fibronectins  
are involved in cell adhesion, cell motility, opsonization, wound  
healing, and maintenance of cell shape.  
CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED  
VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;

```

SQ  SEQUENCE      1256 AA;  137435 MW;  345A4CA0E4D71D9B CRC64;

Query Match      100.0%;  Score 34;  DB 1;  Length 1256;
Best Local Similarity 100.0%;  Pred. No. 22;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      : GRGDSP 6
      |||||
Db      492 GRGDSP 497

RESULT 2
FINC_PLEWA
ID      FINC_PLEWA      STANDARD;      PRT;  1328 AA.
AC      Q91289;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Fibronectin (FN) [Fragment].
OS      Pleurodeles waltlii (Iberian ribbed newt).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC      Pleurodeles.
OX      NCBI_TaxID=8319;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94363379; PubMed=8081872;
RA      Cavalier L., Riou J., Desimone D.W.;
RT      "Amphibian pleurodeles waltl fibronectin: cDNA cloning and
RT      developmental expression of spliced variants.";
RL      Cell Adhes. Commun. 1:83-91(1993).
CC      -1- FUNCTION: Fibronectins bind cell surfaces and various compounds
CC      including collagen, fibrin, heparin, DNA, and actin. Fibronectins
CC      are involved in cell adhesion, cell motility, opsonization, wound
CC      healing, and maintenance of cell shape (By similarity).
CC      -1- SUBUNIT: DIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS,
CC      CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS (BY
CC      SIMILARITY).
CC      -1- SIMILARITY: Contains at least 3 fibronectin type I domains.
CC      -1- SIMILARITY: Contains at least 10 fibronectin type III domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on use
CC      by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X66813; CAA47292.1; -.
DR      HSSP; P02751; 1FNH.
DR      InterPro; IPR006209; EGF like.
DR      InterPro; IPR000083; Fibrinctn1.
DR      InterPro; IPR008957; FN_III-like.
DR      InterPro; IPR003961; FN_III.
DR      InterPro; IPR003962; FnIII_subd.
DR      Pfam; PF00039; fnI; 3.
DR      Pfam; PF00041; fn3; 11.
DR      PRINTS; PR00014; FNTYPEIII.
DR      SMART; SM00058; FN1; 3.
DR      SMART; SM00060; FN3; 9.
DR      PROSITE; PS00022; EGF_1; 1.
DR      PROSITE; PS01253; FIBRONECTIN_1; 1.
KW      Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;
KW      Repeat.
FT      FT      NON TER      1      1
FT      FT      DNA BIND      <1      13
FT      FT      DOMAIN      203      477
FT      FT      DOMAIN      658      929
FT      FT      DOMAIN      1157      1288
FT      FT      DOMAIN      <1      14
FT      FT      DOMAIN      15      110
FT      FT      DOMAIN      111      201
FT      FT      DOMAIN      111      201
      (EXTRA DOMAIN).
      BY SIMILARITY.
      CELL-ATTACHMENT.
      HEPARIN-BINDING 2 (BY SIMILARITY).
      FIBRIN-BINDING 2 (BY SIMILARITY).
      FIBRONECTIN TYPE-III.
      FIBRONECTIN TYPE-III.
      FIBRONECTIN TYPE-III.
      FIBRONECTIN TYPE-III (EXTRA DOMAIN).

```

```

FT DOMAIN 202 292 FIBRONECTIN TYPE-III.
FT DOMAIN 293 383 FIBRONECTIN TYPE-III.
FT DOMAIN 384 477 FIBRONECTIN TYPE-III.
FT DOMAIN 478 567 FIBRONECTIN TYPE-III.
FT DOMAIN 568 657 FIBRONECTIN TYPE-III (EXTRA DOMAIN).
FT DOMAIN 658 749 FIBRONECTIN TYPE-III.
FT DOMAIN 750 838 FIBRONECTIN TYPE-III.
FT DOMAIN 839 929 FIBRONECTIN TYPE-III.
FT DOMAIN 930 1064 CONNECTING STRAND 3 (CS-3) (V REGION).
FT DOMAIN 1065 1143 FIBRONECTIN TYPE-III.
FT DOMAIN 1155 1199 FIBRONECTIN TYPE-I.
FT DOMAIN 1200 1243 FIBRONECTIN TYPE-I.
FT DOMAIN 1245 1287 FIBRONECTIN TYPE-I.
FT DISULFID 1157 1186 BY SIMILARITY.
FT DISULFID 1184 1196 BY SIMILARITY.
FT DISULFID 1202 1229 BY SIMILARITY.
FT DISULFID 1227 1240 BY SIMILARITY.
FT DISULFID 1247 1270 BY SIMILARITY.
FT DISULFID 1268 1284 BY SIMILARITY.
FT DISULFID 1320 1320 INTERCHAIN (WITH C-1324) (BY SIMILARITY).
FT DISULFID 1324 1324 INTERCHAIN (WITH C-1320) (BY SIMILARITY).
FT SITE 461 463 CELL ATTACHMENT SITE.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1328 AA; 145037 MW; E31BF796A1D1E74 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 1328;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels C; Gaps 0;

QY 1 GRGDSF 6
Db 460 GRGDSF 465

RESULT 3
FIND BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT 01-APR-1988 (Rel. 37, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Fibrinectin (FN).
GN FN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_TaxID=9913;
RX MEDLINE=87054047; PubMed=3780752;
RA Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
RT "Complete primary structure of bovine plasma fibronectin.";
RL Eur. J. Biochem. 161:441-453(1986).
RP PARTIAL SEQUENCE.
RX MEDLINE=83117805; PubMed=6218503;
RA Petersen T.E., Thorgerensen H.C., Skorstengaard K., Vibe-Pedersen K.,
RA Sahl P., Sotttrup-Jensens L., Magnusson S.;
RT "Partial primary structure of bovine plasma fibronectin: three types
of internal homology";
RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
RP SEQUENCE OF 2170-2265 FROM N.A.
RX MEDLINE=83221567; PubMed=6304699;
RA Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;
RT "Isolation and characterization of cDNA clones for human and bovine
fibronectins.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
including collagen, fibrin, heparin, DNA, and actin. Fibronectins
are involved in cell adhesion, cell motility, opsonization, wound
healing, and maintenance of cell shape.

```

-!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS; TO A LESSER EXTENT HOMODIMERS.  
 -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 -!- ALTERNATIVE PRODUCTS:  
 Event-Alternative splicing; Named isoforms=1;  
 Comment=A number of isoforms are produced. Each of the "extra domain" and the connecting strand 3 are present in some forms of fibronectin and absent in others;  
 Name=1;  
 IsoId=p07589-1; Sequence=Displayed;  
 -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted by hepatocytes. Cellular FN (dimeric or cross-linked multimeric forms), made by fibroblasts, epithelial and other cell types, is deposited as fibrils in the extracellular matrix.  
 -!- PTM: Sulfated (By similarity).  
 -!- SIMILARITY: Contains 12 fibronectin type I domains.  
 -!- SIMILARITY: Contains 2 fibronectin type II domains.  
 -!- SIMILARITY: Contains 15 fibronectin type III domains.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: K00800; AAA30521.2; -;  
 PIR: A26452; FNBO  
 HSSP: P02751; 2FN2.  
 InterPro: IPR006209; EGF\_like.  
 InterPro: IPR000383; Fibnctnl.  
 InterPro: IPR008957; FN\_III-like.  
 InterPro: IPR003961; FN\_III.  
 InterPro: IPR000562; FN\_Type\_II.  
 InterPro: IPR003962; FNII\_subd.  
 Pfam: PF00039; fn1; 12.  
 Pfam: PF00040; fn2; 2.  
 Pfam: PF00041; fn3; 15.  
 PRINTS: PR00013; ENTPEII.  
 PRINTS: PR00014; ENTPEIII.  
 ProDom: PD000995; FN\_Type\_II; 2.  
 SMART: SMC0058; FN1; 12.  
 SMART: SMC0059; FN2; 2.  
 SMART: SMC0060; FN3; 14.  
 PROSITE: PS00022; EGF\_1; 2.  
 PROSITE: PS00023; FIBRONECTIN\_2; 2.  
 PROSITE: PS01233; FIBRONECTIN\_1; 12.  
 Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;  
 Sulfation; Cell adhesion; Repeat; Alternative splicing;  
 Pyrrolidone carboxylic acid.  
 MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT DOMAIN 21 241 FIBRIN- AND HEPARIN-BINDING 1.  
 FT DOMAIN 277 577 COLLAGEN-BINDING.  
 FT DNA\_BIND 876 1141 CELL-ATTACHMENT.  
 FT DOMAIN 1236 1509 HEPARIN-BINDING 2.  
 FT DOMAIN 1600 1870 FIBRIN-BINDING 2.  
 FT DOMAIN 1991 2216 FIBRONECTIN TYPE-I 1.  
 FT DOMAIN 19 59 FIBRONECTIN TYPE-I 2.  
 FT DOMAIN 64 107 FIBRONECTIN TYPE-I 3.  
 FT DOMAIN 108 151 FIBRONECTIN TYPE-I 4.  
 FT DOMAIN 153 197 FIBRONECTIN TYPE-I 5.  
 FT DOMAIN 198 242 FIBRONECTIN TYPE-I 6.  
 FT DOMAIN 275 314 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 314 373 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 374 438 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 437 480 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 485 527 FIBRONECTIN TYPE-I 8.  
 FT DOMAIN 528 571 FIBRONECTIN TYPE-I 9.  
 FT DOMAIN 578 669 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 688 778 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 779 874 FIBRONECTIN TYPE-III 3.

[illegible]

RP SEQUENCE OF 1441-1548.  
 RX MEDLINE=82265604; PubMed=7050098;  
 RA Piersbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;  
 RT "The cell attachment domain of fibronectin. Determination of the  
 RT primary structure."  
 RL J. Biol. Chem. 257:9593-9597(1982).  
 RN [11]  
 RP SEQUENCE OF 1448-1540 FROM N.A.  
 RX MEDLINE=83290929; PubMed=6688418;  
 RA Oldberg A., Linney E., Ruoslahti E.;  
 RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for  
 RT the cell attachment domain in human fibronectin."  
 RL J. Biol. Chem. 258:10193-10196(1983).  
 RN [12]  
 RP SEQUENCE OF 1448-1540 FROM N.A.  
 RX MEDLINE=86111901; PubMed=3003095;  
 RA Oldberg A., Ruoslahti E.;  
 RT "Evolution of the fibronectin gene. Exon structure of cell attachment  
 RT domain."  
 RL J. Biol. Chem. 261:2113-2116(1986).  
 RN [13]  
 RP SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).  
 RX MEDLINE=85280409; PubMed=292573;  
 RA Berrard M.P., Kolbe M., Weil D., Chu M.-L.;  
 RT "Human cellular fibronectin: comparison of the carboxyl-terminal  
 RT portion with rat identifies primary structural domains separated by  
 RT hypervariable regions."  
 RL Biochemistry 24:2698-2704(1985).  
 RN [14]  
 RP SEQUENCE OF 1712-1739 FROM N.A.  
 RX MEDLINE=87026578; PubMed=3021206;  
 RA Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;  
 RT "Human liver fibronectin complementary DNAs: identification of two  
 RT different messenger RNAs possibly encoding the alpha and beta  
 RT subunits of plasma fibronectin."  
 RL Biochemistry 25:4936-4941(1986).  
 RN [15]  
 RP SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).  
 RC TISSUE=Cartilage;  
 RX MEDLINE=22126816; PubMed=12127832;  
 RA Parker A.E., Boutell J., Carr A., Maciewicz R.A.;  
 RT "Novel cartilage-specific splice variants of fibronectin."  
 RL Osteoarthritis Cartilage 10:528-534(2002).  
 RN [16]  
 RP SEQUENCE OF 32-290.  
 RX MEDLINE=84032463; PubMed=6630202;  
 RA Garcia-Pardo A., Pearlstein E., Frangione B.;  
 RT "Primary structure of human plasma fibronectin. The 29,000-dalton  
 RT NH2-terminal domain."  
 RL J. Biol. Chem. 258:12670-12674(1983).  
 RN [17]  
 RP SEQUENCE OF 369-608, AND COLLAGEN-BINDING.  
 RX MEDLINE=87083265; PubMed=3024962;  
 RA Owens R.J., Baralle F.E.;  
 RT "Mapping the collagen-binding site of human fibronectin by expression  
 RT in Escherichia coli."  
 RL EMBO J. 5:2825-2830(1986).  
 RN [18]  
 RP SULFATION.  
 RX MEDLINE=86042625; PubMed=2414772;  
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
 RT "Tyrosine sulfation of proteins from the human hepatoma cell line  
 RT HepG2."  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
 RN [19]  
 RP O-GLYCOSYLATION OF THR-2064.  
 RX MEDLINE=91190085; PubMed=2012601;  
 RA Tresselt T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,  
 RA Shively J.E., Pande H.;  
 RT "Human plasma fibronectin. Demonstration of structural differences  
 RT between the A- and B-chains in the III CS region."  
 RL Biochem. J. 274:733-738(1991).  
 RN [20]

RP FBLNI-BINDING SITE.  
 RX MEDLINE=93015879; PubMed=1400330;  
 RA Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,  
 RA Argaves W.S.;  
 RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding  
 RT region of fibronectin."  
 RL J. Biol. Chem. 267:20120-20125(1992).  
 RN [21]  
 RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.  
 RX MEDLINE=95081153; PubMed=7989369;  
 RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;  
 RT "Further characterization of the NH2-terminal fibrin-binding site on  
 RT fibronectin."  
 RL J. Biol. Chem. 269:31938-31945(1994).  
 RN [22]  
 RP STRUCTURE BY NMR OF 1447-1540.  
 RX MEDLINE=92162710; PubMed=1311202;  
 RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,  
 RA Campbell I.D.;  
 RT "1H NMR assignment and secondary structure of the cell adhesion type  
 RT III module of fibronectin."  
 RL Biochemistry 31:2068-2073(1992).  
 RN [23]  
 RP STRUCTURE BY NMR OF 1447-1540.  
 RX MEDLINE=93046665; PubMed=1423622;  
 RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;  
 RT "The three-dimensional structure of the tenth type III module of  
 RT fibronectin: an insight into RGD-mediated interactions."  
 RL Cell 71:671-678(1992).  
 RN [24]  
 RP STRUCTURE BY NMR OF 182-275.  
 RX MEDLINE=94141923; PubMed=8308892;  
 RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,  
 RA Campbell I.D.;  
 RT "Solution structure of a pair of fibronectin type 1 modules with  
 RT fibrin binding activity."  
 RL J. Mol. Biol. 235:1302-1311(1994).  
 RN [25]  
 RP STRUCTURE BY NMR OF 32-92.  
 RX MEDLINE=96069779; PubMed=7583666;  
 RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;  
 RT "High-resolution structural studies of the factor XIIIa crosslinking  
 RT site and the first type I module of fibronectin."  
 RL Nat. Struct. Biol. 2:946-950(1995).  
 RN [26]  
 RP STRUCTURE BY NMR OF 406-464.  
 RX MEDLINE=98179538; PubMed=9514732;  
 RA Sticht H., Pickford A.R., Potts J.R., Campbell I.D.;  
 RT "Solution structure of the glycosylated second type 2 module of  
 RT fibronectin."  
 RL J. Mol. Biol. 276:177-187(1998).  
 RN [27]  
 RP STRUCTURE BY NMR OF EXTRA ED-B DOMAIN FROM ISOFORM 7.

Query Match 100.0%; Score 34; DB 1; Length 2386;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
 |||||  
 Db 1523 GRGDSP 1528

RESULT 5  
 FINE MOUSE STANDARD; PRT: 2477 AA.  
 ID FINE MOUSE  
 AC P1276; Q61567; Q61568; Q61569; Q64233; Q80J14;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Fibronectin precursor (FN).  
 GN FNI.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 1-92C FROM N.A.  
RC STRAIN=FVB/N-3; TISSUE=Breast tumor;  
RX MEDLINE=22398257; PubMed=2477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,  
RA Raha S.S., Lequellaro N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE OF 1-28 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=94131313; PubMed=8299972;  
RA Polly P., Nicholas R.C.;  
RT "Sequence of the mouse fibronectin-encoding gene promoter region.";  
RL Gene 137:353-354(1993).  
RN [3]  
RP SEQUENCE OF 562-834 FROM N.A.  
RC STRAIN=NMRI;  
RX MEDLINE=95403556; PubMed=7673336;  
RA Talts J.F., Weller A., Timpl R., Ekblom M., Ekblom P.;  
RT "Regulation of mesenchymal extracellular matrix protein synthesis by  
RT transforming growth factor-beta and glucocorticoids in tumor  
RL stroma.";  
RL J. Cell Sci. 108:2153-2162(1995).  
RN [4]  
RP SEQUENCE OF 899-2376 FROM N.A.  
RA Gorski G., Aros M., Norton P.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 2375-2477 FROM N.A.  
RX MEDLINE=88124987; PubMed=324113;  
RA Blattl S.P., Foster D.N., Ranganathan G., Moses H.L., Getz M.J.;  
RT "Induction of fibronectin gene transcription and mRNA is a primary  
RT response to growth-factor stimulation of AKR-2B cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:1119-1123(1988).  
RN [6]  
RP SEQUENCE OF 2375-2477 FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=930311702; PubMed=1327855;  
RA Khadjian E.W., Salomon C., Leonard N., Tremblay S., Turler H.;  
RT "Fibronectin gene expression in proliferating, quiescent, and SV40-  
RT infected mouse kidney cells.";  
RL Exp. Cell Res. 202:464-470(1992).  
RN [7]  
RP STRUCTURE BY NMR OF 1447-1630.  
RX MEDLINE=98202578; PubMed=9533887;  
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,  
RA Pastor R.W., Krueger S., Zorich D.A.;  
RT "Solution structure and dynamics of linked cell attachment modules of  
RT mouse fibronectin containing the RGD and synergy regions: comparison  
RT with the human fibronectin crystal structure.";  
RL J. Mol. Biol. 277:663-682(1998).  
RN [8]  
RP DOWN-REGULATION BY GLUCOCORTICOIDS.  
RX MEDLINE=21600963; PubMed=11737251;  
RA Gu Y.-C., Talts J.F., Gullberg D., Timpl R., Ekblom M.;  
RT "Glucocorticoids down-regulate the extracellular matrix proteins  
RT fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";  
RL Eur. J. Haematol. 67:176-184(2001).  
RN [9]  
RP FUNCTION: Fibronectins bind cell surfaces and various compounds  
RC including collagen, fibrin, heparin, DNA, and actin. Fibronectins  
RX are involved in cell adhesion, cell motility, opsonization, wound  
RY healing, and maintenance of cell shape.  
RN [10]  
RP SUBMIT: Mostly heterodimers or multimers of alternatively spliced  
RC variants, connected by 2 disulfide bonds near the carboxyl ends;  
RX to a lesser extend homodimers. Interacts with FBLN1 (By  
RY similarity).  
RN [11]  
RP SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
RX ALTERNATIVE PRODUCTS:  
RY Event=Alternative splicing; Named isoforms=1;  
RY Comment=A number of isoforms are produced. Each of the "extra  
RY domain" and the connecting strand 3 are present in some forms of  
RY fibronectin and absent in others;  
RY Names1;  
RY IsoId=P11276-1; Sequence=Displayed;  
RX TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted  
RY by hepatocytes. Cellular FN (dimeric or cross-linked multimeric  
RY forms), made by fibroblasts, epithelial and other cell types, is  
RY deposited as fibrils in the extracellular matrix.  
RX INDUCTION: Glucocorticoids suppressed mRNA expression and protein  
RY synthesis.  
RY PTM: Sulfated (By similarity).  
RX SIMILARITY: Contains 12 fibronectin type I domains.  
RY SIMILARITY: Contains 2 fibronectin type II domains.  
RX SIMILARITY: Contains 17 fibronectin type III domains.  
RN [12]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RY between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RY the European Bioinformatics Institute. There are no restrictions on its  
RY use by non-profit institutions as long as its content is in no way  
RY modified and this statement is not removed. Usage by and for commercial  
RY entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RY or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RX EMBL; BC051082; AAHS1082.1; -  
RY EMBL; Z22729; CAA80422.1; -  
RX EMBL; X82402; CAA57796.1; -  
RY EMBL; X93167; CAA63654.1; -  
RX EMBL; M18194; AAA37636.1; -  
RY EMBL; S45860; AAB23491.1; -  
RX PIR; A49173; A49173.  
RY PIR; I48349; I48349.  
RX PDB; 1MEN; 29-APR-98.  
RY PDB; 2MEN; 29-APR-98.  
RX MGD; MGI-95566; Fnl.  
RY GO; GO:0007155; P:cell adhesion; IDA.  
RX InterPro; IPR006209; EGF-like.  
RY InterPro; IPR000083; Fibrnctnl.  
RX InterPro; IPR008957; FN III-like.  
RY InterPro; IPR003961; FN III.  
RX InterPro; IPR000562; FN\_Type\_II.  
RY InterPro; IPR008924; FNIII\_subd.  
RX PRINTS; PR00014; FNTYPEIII.  
RY SMART; SM00059; FN1; 4.  
RX SMART; SM00060; FN3; 12.  
RY PROSITE; PS00022; EGF\_1; 2.  
RX PROSITE; PS00023; FIBRONECTIN\_2; 2.  
RY PROSITE; PS01253; FIBRONECTIN\_1; 12.  
RX Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;  
RY Sulfation; Cell adhesion; Repeat; Alternative splicing; Signal;  
RY 3D-structure. 1 32 BY SIMILARITY.  
RX SIGNAL 33 2477 FIBRONECTIN.  
RY CHAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.  
RX DOMAIN 308 608 COLLAGEN-BINDING.  
RY DNA\_BIND 906 1171





DR InterPro; IPR000562; FN\_Type II.  
 DR InterPro; IPR003962; FNIII\_subd.  
 DR Pfam; PF00039; fn1; 12.  
 DR Pfam; PF00040; fn2; 2.  
 DR Pfam; PF00041; fn3; 17.  
 DR PRINTS; PR00013; FNTYPEII.  
 DR PRINTS; PR00014; FNTYPEIII.  
 DR ProDom; PD000995; FN\_Type\_II; 2.  
 DR SMART; SM00058; FN1; 12.  
 DR SMART; SM00059; FN2; 2.  
 DR SMART; SM00060; FN3; 13.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 12.  
 DR Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;  
 KW Sulfatation; Cell adhesion; Repeat; Alternative splicing; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 2477 FIBRONECTIN.  
 FT DOMAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.  
 FT DOMAIN 308 608 COLLAGEN-BINDING.  
 FT DNA\_BIND 906 1171  
 FT DOMAIN 1357 1630  
 FT DOMAIN 1811 2081  
 FT DOMAIN 2296 2427  
 FT DOMAIN 51 91  
 FT DOMAIN 96 139 FIBRONECTIN TYPE-I 1.  
 FT DOMAIN 140 183 FIBRONECTIN TYPE-I 2.  
 FT DOMAIN 185 229 FIBRONECTIN TYPE-I 3.  
 FT DOMAIN 230 274 FIBRONECTIN TYPE-I 4.  
 FT DOMAIN 306 345 FIBRONECTIN TYPE-I 5.  
 FT DOMAIN 345 404 FIBRONECTIN TYPE-I 6.  
 FT DOMAIN 405 469 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 468 511 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 516 558 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 559 602 FIBRONECTIN TYPE-I 8.  
 FT DOMAIN 609 706 FIBRONECTIN TYPE-I 9.  
 FT DOMAIN 707 808 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 809 903 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 904 994 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 995 1084 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 1085 1172 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 1173 1264 FIBRONECTIN TYPE-III 4.  
 FT DOMAIN 1265 1355 FIBRONECTIN TYPE-III 5.  
 FT DOMAIN 1356 1446 FIBRONECTIN TYPE-III 6.  
 FT DOMAIN 1447 1536 FIBRONECTIN TYPE-III 7.  
 FT DOMAIN 1537 1630 FIBRONECTIN TYPE-III 8.  
 FT DOMAIN 1631 1720 FIBRONECTIN TYPE-III 9.  
 FT DOMAIN 1721 1810 FIBRONECTIN TYPE-III 10.  
 FT DOMAIN 1811 1902 FIBRONECTIN TYPE-III 11.  
 FT DOMAIN 1903 1991 FIBRONECTIN TYPE-III 12.  
 FT DOMAIN 1992 2081 FIBRONECTIN TYPE-III 13.  
 FT DOMAIN 2082 2201 FIBRONECTIN TYPE-III 14.  
 FT DOMAIN 2202 2283 FIBRONECTIN TYPE-III 15.  
 FT DOMAIN 2294 2338 FIBRONECTIN TYPE-III 16.  
 FT DOMAIN 2339 2381 FIBRONECTIN TYPE-III 17.  
 FT DOMAIN 2383 2426 FIBRONECTIN TYPE-I 10.  
 FT SITE 1614 1616 FIBRONECTIN TYPE-I 11.  
 FT SITE 2181 2183 FIBRONECTIN TYPE-I 12.  
 FT DISULFID 53 79 CELL ATTACHMENT SITE.  
 FT DISULFID 77 88 CELL ATTACHMENT SITE.  
 FT DISULFID 98 126 BY SIMILARITY.  
 FT DISULFID 124 136 BY SIMILARITY.  
 FT DISULFID 142 170 BY SIMILARITY.  
 FT DISULFID 168 180 BY SIMILARITY.  
 FT DISULFID 187 216 BY SIMILARITY.  
 FT DISULFID 214 226 BY SIMILARITY.  
 FT DISULFID 232 261 BY SIMILARITY.  
 FT DISULFID 259 271 BY SIMILARITY.  
 FT DISULFID 308 335 BY SIMILARITY.  
 FT DISULFID 333 342 BY SIMILARITY.  
 FT DISULFID 360 386 BY SIMILARITY.  
 FT DISULFID 374 401 BY SIMILARITY.  
 FT DISULFID 420 446 BY SIMILARITY.

FT DISULFID 434 461 BY SIMILARITY.  
 FT DISULFID 470 498 BY SIMILARITY.  
 FT DISULFID 496 508 BY SIMILARITY.  
 FT DISULFID 518 545 BY SIMILARITY.  
 FT DISULFID 543 555 BY SIMILARITY.  
 FT DISULFID 561 589 BY SIMILARITY.  
 FT DISULFID 587 599 BY SIMILARITY.  
 FT DISULFID 2296 2325 BY SIMILARITY.  
 FT DISULFID 2323 2335 BY SIMILARITY.  
 FT DISULFID 2341 2368 BY SIMILARITY.  
 FT DISULFID 2366 2378 BY SIMILARITY.  
 FT DISULFID 2385 2409 BY SIMILARITY.  
 FT DISULFID 2407 2423 BY SIMILARITY.  
 FT DISULFID 2458 2458 INTERCHAIN (WITH C-2462).  
 FT DISULFID 2462 2462 INTERCHAIN (WITH C-2458).  
 FT MOD\_RES 875 880 SULEFATION (POTENTIAL).  
 FT MOD\_RES 880 880 SULEFATION (POTENTIAL).  
 FT MOD\_RES 2392 2392 SULEFATION (POTENTIAL).  
 FT MOD\_RES 2475 2475 PHOSPHORYLATION (BY SIMILARITY).  
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 876 876 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1006 1006 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 2198 2198 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 2154 2154 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 FT VARSPLIC 1720 1809 Missing (in isoform 2).  
 FT VARSPLIC 2082 2106 /FTid=VSP\_003258.  
 FT VARSPLIC 2082 2200 /FTid=VSP\_003259.  
 FT CONFLICT 2318 2318 Missing (in isoform 4).  
 FT CONFLICT 2318 2318 /FTid=VSP\_003260.  
 FT CONFLICT 2318 2318 G -> A (IN REF. 3).  
 SQ SEQUENCE 2477 AA; 272510 MW; B4391A472CEDEB5 CRC64;  
 Query Match 100.0%; Score 34; DB 1; Length 2477;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 Db 1613 GRGDSP 1618  
 RESULT 7  
 FINE\_XENLA STANDARD; PRT; 2481 AA.  
 ID FINE\_XENLA STANDARD; PRT; 2481 AA.  
 AC 091740;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Fibrinectin precursor.  
 GN FN1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92111942; PubMed=1730390;  
 RA Desimone D.W., Norton P.A., Hynes R.O.;  
 RT "Identification and characterization of alternatively spliced  
 RL fibrinectin mRNAs expressed in early Xenopus embryos.";  
 RL Dev. Biol. 149:357-369(1992).  
 CC -!- FUNCTION: Fibrinectins bind cell surfaces and various compounds  
 CC including collagen, fibrin, heparin, DNA and actin. Fibrinectins  
 CC are involved in cell adhesion, cell motility, opsonization, wound  
 CC healing, and maintenance of cell shape (By similarity).  
 CC -!- SUBUNIT: DIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS,  
 CC CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS (BY  
 CC SIMILARITY).

CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=1;  
 CC Comment=A number of isoforms are produced. Each of the 'extra  
 CC domain' and the connecting strand 3 are present in some forms of  
 CC fibronectin and absent in others;  
 CC Name=1;  
 CC IsoId=C91740-1; Sequence=Displayed;  
 CC -!- TISSUE SPECIFICITY: IN EARLY XENOPUS EMBRYO, CELLULAR FORMS OF  
 CC FIBRONECTIN PREDOMINATE WHICH INCLUDE BOTH EXTRA DOMAINS. IN  
 CC FIBRONECTIN OF EMBRYONIC AND ADULT LIVER THE CONNECTING STRAND 3  
 CC CAN BE EITHER COMPLETELY EXCLUDED OR INCLUDED.  
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.  
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.  
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isdb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; M77820; AAA49707.1; -;  
 CC HSSP; P02751; 2FN2  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR000083; Fibrctnln.  
 CC InterPro; IPR008957; FN III-like.  
 CC InterPro; IPR003961; FN III.  
 CC InterPro; IPR000562; FN\_Type\_II.  
 CC Pfam; PF00039; fn1; 12.  
 CC Pfam; PF00040; fn2; 2.  
 CC Pfam; PF00041; fn3; 17.  
 CC PRINTS; PR00013; FNYPEII.  
 CC ProDom; PD000995; FN\_Type\_II; 2.  
 CC SMART; SM00058; FN1\_12.  
 CC SMART; SM00059; FN2; 2.  
 CC SMART; SM00060; FN3; 17.  
 CC PROSITE; PS00022; EGF\_1; 2.  
 CC PROSITE; PS01253; FIBRONECTIN 1; 11.  
 CC Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;  
 CC Repeat; Alternative splicing; Signal.  
 CC SIGNAL 31  
 CC CHAIN 32 2481 FIBRONECTIN.  
 CC DOMAIN 55 275 FIBRIN- AND HEPARIN-BINDING 1.  
 CC DOMAIN 309 609 COLLAGEN-BINDING.  
 CC DNA BIND 907 1172 BY SIMILARITY.  
 CC DOMAIN 1358 1631 CELL-ATTACHMENT.  
 CC DOMAIN 1812 2082 HEPARIN-BINDING 2.  
 CC DOMAIN 2301 2432 FIBRIN-BINDING 2.  
 CC DOMAIN 53 93 FIBRONECTIN TYPE-I 1.  
 CC DOMAIN 98 141 FIBRONECTIN TYPE-I 2.  
 CC DOMAIN 142 185 FIBRONECTIN TYPE-I 3.  
 CC DOMAIN 187 231 FIBRONECTIN TYPE-I 4.  
 CC DOMAIN 232 276 FIBRONECTIN TYPE-I 5.  
 CC DOMAIN 307 346 FIBRONECTIN TYPE-I 6.  
 CC DOMAIN 346 405 FIBRONECTIN TYPE-II 1.  
 CC DOMAIN 406 470 FIBRONECTIN TYPE-II 2.  
 CC DOMAIN 469 512 FIBRONECTIN TYPE- 7.  
 CC DOMAIN 517 559 FIBRONECTIN TYPE-I 8.  
 CC DOMAIN 560 603 FIBRONECTIN TYPE-I 9.  
 CC DOMAIN 610 707 FIBRONECTIN TYPE-III 1.  
 CC DOMAIN 708 809 FIBRONECTIN TYPE-III 2.  
 CC DOMAIN 810 904 FIBRONECTIN TYPE-III 3.  
 CC DOMAIN 905 995 FIBRONECTIN TYPE-III 4.  
 CC DOMAIN 996 1085 FIBRONECTIN TYPE-III 5.  
 CC DOMAIN 1086 1173 FIBRONECTIN TYPE-III 6.  
 CC DOMAIN 1174 1265 FIBRONECTIN TYPE-III 7.  
 CC DOMAIN 1266 1356 FIBRONECTIN TYPE-III 8 (EXTRA DOMAIN).  
 CC DOMAIN 1357 1447 FIBRONECTIN TYPE-III 9.  
 CC DOMAIN 1448 1537 FIBRONECTIN TYPE-III 10.  
 CC DOMAIN 1538 1631 FIBRONECTIN TYPE-III 11.  
 CC DOMAIN 1632 1721 FIBRONECTIN TYPE-III 12.

FT DOMAIN 1722 1811 FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN).  
 FT DOMAIN 1812 1903 FIBRONECTIN TYPE-III 14.  
 FT DOMAIN 1904 1992 FIBRONECTIN TYPE-III 15.  
 FT DOMAIN 1993 2082 FIBRONECTIN TYPE-III 16.  
 FT DOMAIN 2083 2205 CONNECTING STRAND 3 (CS-3) (V REGION).  
 FT DOMAIN 2206 2287 FIBRONECTIN TYPE-III 17.  
 FT DOMAIN 2288 2344 FIBRONECTIN TYPE-I 10.  
 FT DOMAIN 2345 2386 FIBRONECTIN TYPE-I 11.  
 FT DOMAIN 2387 2431 FIBRONECTIN TYPE-I 12.  
 FT SITE 1615 1617 CELL ATTACHMENT SITE.  
 FT DISULFID 55 81 BY SIMILARITY.  
 FT DISULFID 79 90 BY SIMILARITY.  
 FT DISULFID 100 128 BY SIMILARITY.  
 FT DISULFID 126 138 BY SIMILARITY.  
 FT DISULFID 144 172 BY SIMILARITY.  
 FT DISULFID 170 182 BY SIMILARITY.  
 FT DISULFID 189 218 BY SIMILARITY.  
 FT DISULFID 216 228 BY SIMILARITY.  
 FT DISULFID 234 263 BY SIMILARITY.  
 FT DISULFID 261 273 BY SIMILARITY.  
 FT DISULFID 309 336 BY SIMILARITY.  
 FT DISULFID 334 343 BY SIMILARITY.  
 FT DISULFID 361 387 BY SIMILARITY.  
 FT DISULFID 375 402 BY SIMILARITY.  
 FT DISULFID 421 447 BY SIMILARITY.  
 FT DISULFID 471 499 BY SIMILARITY.  
 FT DISULFID 497 509 BY SIMILARITY.  
 FT DISULFID 519 546 BY SIMILARITY.  
 FT DISULFID 544 556 BY SIMILARITY.  
 FT DISULFID 562 590 BY SIMILARITY.  
 FT DISULFID 588 600 BY SIMILARITY.  
 FT DISULFID 2301 2330 BY SIMILARITY.  
 FT DISULFID 2328 2340 BY SIMILARITY.  
 FT DISULFID 2346 2373 BY SIMILARITY.  
 FT DISULFID 2371 2383 BY SIMILARITY.  
 FT DISULFID 2390 2414 BY SIMILARITY.  
 FT DISULFID 2412 2428 BY SIMILARITY.  
 FT DISULFID 2459 2459 INTERCHAIN (WITH C-2463) (BY SIMILARITY).  
 FT DISULFID 2463 2463 INTERCHAIN (WITH C-2459) (BY SIMILARITY).  
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 543 543 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 877 877 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1244 1244 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1291 1291 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 2202 2202 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 2481 AA; 272678 MW; 7E47DF4F6CE72C93 CRC64;  
 Query Match 100.0%; Score 34; DB 1; Length 2481;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 Db 1614 GRGDSP 1619  
 RESULT 8  
 HUPK RHILV  
 ID HUPK RHILV STANDARD; PRT; 370 AA.  
 AC 28153;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Hydrogenase expression/formation protein hupk.  
 GN HUPK.  
 OS Rhizobium leguminosarum (biovar viciae).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=128C53;





```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L18880; AAB96843.1; -
CC EMBL; L13300; ABA40557.1; -
CC EMBL; L13299; ABA40557.1; JOINED.
CC PIR; T10108; T10108.
CC MGD; MGI; 98927; Vc-.
CC GO; GO:0005912; C:adherens junction; IDA.
CC GO; GO:0005925; C:focal adhesion; IDA.
CC GO; GO:0025515; F:protein binding; IPI.
CC GO; GO:0037155; P:cell adhesion; IDA.
CC GO; GO:0030032; P:lamellipodium formation; IDA.
CC GO; GO:0030334; P:regulation of cell migration; IDA.
CC InterPro; IPR006077; Vinculin/catenin.
CC InterPro; IPR006633; Vinculin_2.
CC Pfam; PF01044; Vinculin; 1.
CC PRINTS; PR00806; VINCULIN.
CC PRODOM; PD02589; Vinculin_2; 1.
CC PROSITE; PS00663; VINCULIN_1; 1.
CC PROSITE; PS00664; VINCULIN_2; 3.
CC Cell adhesion; Actin-binding; Cytoskeleton; Structural protein;
KW Phosphorylation; Repeat; Lipoprotein.
FT INIT_MET 0
FT DOMAIN 167 207
FT DOMAIN 258 588
FT REPEAT 258 368
FT REPEAT 369 478
FT REPEAT 479 588
FT DOMAIN 836 977
FT MOD_RES 536 536
FT MOD_RES 821 821
SQ SEQUENCE 1065 AA; 116674 MW; 4DAB5899024CCDC1 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 1065;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6
Db 451 GKGDSP 456

RESULT 12
VINC HUMAN
ID VINC HUMAN STANDARD; PRT; 1133 AA.
AC P18206; Q16450;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 42; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Vinculin (Metavinculin).
GN VCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Endothelial cells;
RX MEDLINE=903332642; PubMed=2116004;
RA Weller P.A., Ogryzko E.P., Corben E.B., Zhidkova N.I., Patel B.,
RA Price G.J., Spurr N.K., Kotliansky V.E., Critchley D.R.;
RT "Complete sequence of human vinculin and assignment of the gene to
RT chromosome 10.";
RJ Proc. Natl. Acad. Sci. U.S.A. 87:5667-5671(1990).
RN [2]
RP SEQUENCE OF 1-55 FROM N.A.
RX MEDLINE=93179440; PubMed=8440716;
RA Moiseyeva E.P., Weller P.A., Zhidkova N.I., Corben E.B., Patel B.,
RA Jasinska I., Kotliansky V.E., Critchley D.R.;
RT "Organization of the human gene encoding the cytoskeletal protein
RT vinculin and the sequence of the vinculin promoter.";
RL J. Biol. Chem. 268:4318-4325(1993).

```

```

RN [3]
RP SEQUENCE OF 853-1050 FROM N.A. (ISOFORM 2).
RC TISSUE=Uterus;
RX MEDLINE=92174935; PubMed=1339348;
RA Kotliansky V.E., Ogryzko E.P., Zhidkova N.I., Weller P.A.,
RA Critchley D.R., Vancompernelle K., Vandekerckhove J., Strasser P.,
RA Way M., Gimona M., Small J.V.;
RT "An additional exon in the human vinculin gene specifically encodes
RT meta-vinculin-specific difference peptide. Cross-species comparison
RT reveals variable and conserved motifs in the meta-vinculin insert.";
RL Eur. J. Biochem. 204:767-772(1992).
CC -!- FUNCTION: Involved in cell adhesion. May be involved in the
CC attachment of the actin-based microfilaments to the p-asma
CC membrane.
CC -!- SUBUNIT: Exhibits self-association properties.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic face of adhesion plaques.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2; Synonyms=Metavinculin;
CC IsoId=PI8206-1; Sequence=Displayed;
CC Note=Differs from vinculin by the insertion of a 68 residues
CC domain near the C-terminus;
CC Name=1; Synonyms=Vinculin;
CC IsoId=PI8206-2; Sequence=VSP 006731;
CC -!- TISSUE SPECIFICITY: Metavinculin is muscle-specific.
CC -!- PTM: Phosphorylated on serines, threonines and tyrosines and
CC acylated by myristic acid and/or palmitic acid.
CC -!- SIMILARITY: STRONG, TO OTHER VINCULINS AND TO ALPHA-CATENINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13308; ABA61283.1; -
CC EMBL; L04933; ABA61271.1; -
CC EMBL; S87180; AAB21656.1; -
CC EMBL; S87175; AAB21656.1; JOINED.
CC EMBL; S87178; AAB21656.1; JOINED.
CC EMBL; S87223; AAB21657.1; -
CC EMBL; S87218; AAB21657.1; JOINED.
CC PIR; A35955; A35955.
CC SWISS-2DPAGE; P18206; HUMAN.
CC Aarhus/Ghent-2DPAGE; 3708; IEF.
CC HSC-2DPAGE; P18206; HUMAN.
CC Genew; HGNC:12665; VCL.
CC MIM; 193065; -
CC InterPro; IPR006077; Vinculin/catenin.
CC InterPro; IPR006633; Vinculin_2.
CC Pfam; PF01044; Vinculin; 1.
CC PRINTS; PR00806; VINCULIN.
CC PRODOM; PD02589; Vinculin_2; 1.
CC PROSITE; PS00663; VINCULIN_1; 1.
CC PROSITE; PS00664; VINCULIN_2; 3.
CC Cell adhesion; Actin-binding; Cytoskeleton; Structural protein;
KW Phosphorylation; Repeat; Alternative splicing; Lipoprotein.
FT INIT_MET 0
FT DOMAIN 167 207
FT DOMAIN 258 588
FT REPEAT 258 368
FT REPEAT 369 478
FT REPEAT 479 588
FT DOMAIN 836 877
FT MOD_RES 536 536
FT MOD_RES 821 821
FT VARSPPLIC 915 982
SQ SEQUENCE 1133 AA; 123660 MW; 80C9E1260EC25879 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 1133;

```

Best Local Similarity 83.3%; Pred. No. 92;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
|:|||||  
Db 451 GKGDSP 456

RESULT 13  
S3B1 HUMAN STANDARD; PRT; 1304 AA.  
ID S3B1 HUMAN AC 075533;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Splicing factor 3B subunit 1 (Spliceosome associated protein; 55) (SAP  
DE 155) (SF3B155) (Pre-mRNA splicing factor SF3B 155 kDa subunit).  
GN SF3B1 OR SAPI55.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.  
EX MEDLINE=98252426; PubMed=9585501;  
RA Wang C., Chua K., Seghezzi W., Lees E., Gozani O., Reed R.;  
RT "Phosphorylation of spliceosomal protein SAP 155 coupled with splicing  
RT catalysis."  
RL Genes Dev. 12:1409-1414(1998).  
RN [2]  
RP SEQUENCE OF 1011-1304 FROM N.A.  
RC TISSUE=Brain;  
RA Yu W., Gibbs R.A.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP CHARACTERIZATION OF THE SPICEOSOME.  
EX MEDLINE=20337962; PubMed=10882114;  
RA Das R., Zhou Z., Reed R.;  
RT "Functional association of U2 snRNP with the ATP-independent  
RT spliceosomal complex B."  
RL Mol. Cell 5:779-787(2000).  
RN [4]  
CC -!- FUNCTION: Subunit of the splicing factor SF3B required for 'A'  
CC complex assembly formed by the stable binding of U2 snRNP to the  
CC branchpoint sequence (bps) in pre-mRNA. Sequence independent  
CC binding of SF3A/SF3B complex upstream of the branch site is  
CC essential, it may anchor U2 snRNP to the pre-mRNA. May also be  
CC involved in the assembly of the 'E' complex. Belongs also to the  
CC minor U12-dependent spliceosome, which is involved in the splicing  
CC of rare class of nuclear pre-mRNA intron.  
CC -!- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF  
CC FOUR SUBUNITS; SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,  
CC SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A  
CC 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS  
CC COMPLEX (U2 SNRNP). SF3B1 INTERACTS DIRECTLY WITH THE SPLICING  
CC FACTOR U2AF.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR SPECKLES. DURING MITOSIS,  
CC TRANSIENTLY DISPERSED FROM THE NUCLEAR SPECKLES TO THE CYTOPLASM.  
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION OCCURS CONCOMITANT WITH THE  
CC CATALYTIC STEPS OF SPLICING.  
CC -!- SIMILARITY: Belongs to the SF3B1 family.  
CC -!- SIMILARITY: Contains ? HEAT repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF054284; AAC97189.1; -  
CC EMBL; AF070540; AAC28633.1; -  
CC Genew; HGNC:10769; SF3B1.  
CC

DR GK; 075533; -;  
DR MM; 605590; -;  
DR GO; GO:0005681; C:spliceosome complex; NAS.  
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; NAS.  
DR GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; NAS.  
DR InterPro; IPR008938; ARM.  
KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;  
KW Phosphorylation; Repeat.  
SQ SEQUENCE 1304 AA; 145814 MW; 40DB21A620915A7 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 1304;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
|:|||||  
Db 269 GRGDSP 274

RESULT 14  
S3B1 MOUSE STANDARD; PRT; 1304 AA.  
ID S3B1 MOUSE AC Q99NE9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Splicing factor 3B subunit 1 (Spliceosome associated protein 155) (SAP  
DE 155) (SF3B155) (Pre-mRNA splicing factor SF3B 155 kDa subunit).  
GN SF3B1 OR SAPI55.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
EX MEDLINE=21189985; PubMed=11252167;  
RA Iacono K., Abe K., Tomaru Y., Okazaki Y., Hayashizaki Y., Koseki H.;  
RT "Molecular cloning, genetic mapping, and expression of the mouse Sf3b1  
RT (SAPI55) gene for the U2snRNP component of spliceosome."  
RL Mamm. Genome 12:192-198(2001).  
RN [2]  
CC -!- FUNCTION: Subunit of the splicing factor SF3B required for 'A'  
CC complex assembly formed by the stable binding of U2 snRNP to the  
CC branchpoint sequence (bps) in pre-mRNA. Sequence independent  
CC binding of SF3A/SF3B complex upstream of the branch site is  
CC essential, it may anchor U2 snRNP to the pre-mRNA. May also be  
CC involved in the assembly of the 'E' complex. Belongs also to the  
CC minor U12-dependent spliceosome, which is involved in the splicing  
CC of rare class of nuclear pre-mRNA intron (By similarity).  
CC -!- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF  
CC FOUR SUBUNITS; SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,  
CC SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A  
CC 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS  
CC COMPLEX (U2 SNRNP). SF3B1 INTERACTS DIRECTLY WITH THE SPLICING  
CC FACTOR U2AF (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- TISSUE SPECIFICITY: Ubiquitous.  
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION OCCURS CONCOMITANT WITH THE  
CC CATALYTIC STEPS OF SPLICING (BY SIMILARITY).  
CC -!- SIMILARITY: Belongs to the SF3B1 family.  
CC -!- SIMILARITY: Contains ? HEAT repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB037890; BAB40140.1; -  
CC MGD; MGI:1922339; Sf3b1.  
CC InterPro; IPR008938; ARM.  
CC

KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;  
KW Phosphorylation; Repeat.  
SQ SEQUENCE 1304 AA; 145816 MW; 12F051757D2A2DEE CRC64;

Query Match 91.2%; Score 31; DB 1; Length 1304;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6

Db 269 GRGDTF 274

Db 272 GRGDTF 277

Search completed: April 16, 2004, 07:22:14  
Job time : 11.5 secs

RESULT 15  
S3B1\_XENLA  
ID S3B1\_XENLA STANDARD; PRT; 1307 AA.  
AC OS7683;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Splicing factor 3B subunit 1 (Spliceosome associated protein 155) (SAP 155) (SF3B155) (Pre-mRNA splicing factor SF3B 155 kDa subunit) (146 kDa nuclear protein).  
DE XENOPUS laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=83355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98099695; PubMed=9436997;  
RA Schmidt-Zachmann M.S., Knecht S., Kraemer A.;  
RT "Molecular characterization of a novel, widespread nuclear protein that colocalizes with spliceosome components.";  
RL Mol. Biol. Cell 9:143-160(1998).  
CC -!- FUNCTION: Subunit of the splicing factor SF3B required for 'A' complex assembly formed by the stable binding of U2 snRNP to the branchpoint sequence (bps) in pre-mRNA. Sequence independent binding of SF3A/SF3B complex upstream of the branch site is essential, it may anchor U2 snRNP to the pre-mRNA. May also be involved in the assembly of the 'E' complex. Belongs also to the minor U12-dependent spliceosome, which is involved in the splicing of rare class of nuclear pre-mRNA intron (By similarity).  
CC -!- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF FOUR SUBUNITS: SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145, SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEIN COMPLEX (U2 snRNP) (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: Belongs to the SF3B1 family.  
CC -!- SIMILARITY: Contains ? HEAT repeats.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; Y08997; CAA70201.1; -.  
DR PIR; T30887; T30887.  
DR InterPro; IPR008338; ARM.  
KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;  
KW Repeat.  
SQ SEQUENCE 1307 AA; 146213 MW; 6B226D5F036C69B3 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 1307;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6

|||||

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 47.4 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-7  
Perfect score: 34  
Sequence: 1 GRGDSP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	23	4 Q9UC00	Q9UC00 homo sapien
2	34	100.0	83	16 Q7WGA3	Q7WGA3 bordetella
3	34	100.0	83	16 Q7W4S7	Q7W4S7 bordetella
4	34	100.0	83	16 Q7VU91	Q7VU91 bordetella
5	34	100.0	86	2 Q9E2D2	Q9E2D2 rhizobium m
6	34	100.0	208	16 Q9XSG0	Q9XSG0 oryctolagus
7	34	100.0	293	6 Q8CCB5	Q8CCB5 mus musculus
8	34	100.0	363	11 Q9HQN2	Q9HQN2 halobacteri
9	34	100.0	441	17 Q8NC49	Q8NC49 homo sapien
10	34	100.0	562	4 Q8TYQ5	Q8TYQ5 methanopyru
11	34	100.0	618	17 Q8TYQ5	Q8TYQ5 arabisidopsis
12	34	100.0	659	10 Q8GWX7	Q8GWX7 arabisidopsis
13	34	100.0	691	10 Q9M334	Q9M334 arabisidopsis
14	34	100.0	2478	13 Q93406	Q93406 brachydanio
15	34	100.0	3173	16 Q882M6	Q882M6 pseudomonas
16	31	91.2	76	16 Q7UUR4	Q7UUR4 rhodopirell

17	31	91.2	86	10 Q7XVE5	Q7XVE5 oryza sativ
18	31	91.2	144	11 Q8C3P1	Q8C3P1 mus musculus
19	31	91.2	161	11 Q9CVT7	Q9CVT7 mus musculus
20	31	91.2	247	3 Q3CIR3	Q3CIR3 emericella
21	31	91.2	250	16 Q8XXJ6	Q8XXJ6 ralistonia s
22	31	91.2	260	16 Q92NT1	Q92NT1 rhizobium m
23	31	91.2	328	16 Q88P14	Q88P14 pseudomonas
24	31	91.2	347	16 Q8FR63	Q8FR63 corynebacte
25	31	91.2	354	5 Q9BLU0	Q9BLU0 bombyx mori
26	31	91.2	450	16 Q9AA55	Q9AA55 caulobacter
27	31	91.2	477	4 Q9YGR9	Q9YGR9 homo sapien
28	31	91.2	490	11 Q9CSK5	Q9CSK5 mus musculus
29	31	91.2	496	11 Q9ET34	Q9ET34 rattus norv
30	31	91.2	500	13 Q7ZFA4	Q7ZFA4 xenopus lae
31	31	91.2	507	4 Q96T04	Q96T04 homo sapien
32	31	91.2	548	16 Q8FP12	Q8FP12 corynebacte
33	31	91.2	562	4 Q96SV4	Q96SV4 homo sapien
34	31	91.2	574	4 Q96SM8	Q96SM8 homo sapien
35	31	91.2	575	16 Q89P25	Q89P25 bradyrhizob
36	31	91.2	578	16 Q7UIK9	Q7UIK9 rhodopirell
37	31	91.2	667	2 Q9APD3	Q9APD3 thiobacilli
38	31	91.2	699	4 Q96SW4	Q96SW4 homo sapien
39	31	91.2	711	11 Q88U81	Q88U81 mus musculus
40	31	91.2	789	4 Q7Z497	Q7Z497 homo sapien
41	31	91.2	1005	11 Q9EQ71	Q9EQ71 mus musculus
42	31	91.2	1011	5 Q96J38	Q96J38 plasmodium
43	31	91.2	1066	4 Q8IXU7	Q8IXU7 homo sapien
44	31	91.2	1066	11 Q922D9	Q922D9 mus musculus
45	31	91.2	1066	11 Q8BP32	Q8BP32 mus musculus

ALIGNMENTS

RESULT 1

Q9UC00 PRELIMINARY; PRT; 23 AA.  
AC Q9UC00; 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DE 01-JUN-2003 (TREMREL. 24, Last annotation update)  
EN Enhancement of wound HEALING process.  
OS Homo sapiens (Human)  
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=95130623; PubMed=7829572;  
RA Pierschbacher M.D., Polarek J.W., Craig W.S., Tschopp J.F.,  
RA Sipes N.J., Harper J.R.;  
RL J. Cell. Biochem. 56:150-154(1994).  
DR GO; GO:0009611; P:response to wounding; TAS.  
SQ SEQUENCE 23 AA; 2268 MW; CE73999CB9933891 CRC64;

Query Match 100.0%; Score 34; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 4.4; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

Qy 1 GRGDSP 6

Db 1 GRGDSP 6

RESULT 2

Q7WGA3 PRELIMINARY; PRT; 83 AA.  
AC Q7WGA3; 01-OCT-2003 (TREMREL. 25, Created)  
DT 01-OCT-2003 (TREMREL. 25, Last sequence update)  
DE 01-OCT-2003 (TREMREL. 25, Last annotation update)  
EN Putative cell division topological specificity factor.  
GN MINE OR BB4016.



OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus).
CC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
CC	Alcaligenaceae; Bordetella.
CX	NCBI_TaxID=518;
RN	[...]
RP	SEQUENCE FROM N.A.
RC	STRAIN=RB50 / ATCC BAA-588;
RX	MEDLINE=22827954; PubMed=12910271;
RA	Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA	Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA	Leather S., Moule S., Norberczak E., O'Neill S., Ormond D., Price C.,
RA	Rabinowitz E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,
RT	Bordetella parapertussis and Bordetella bronchiseptica.";
RL	Nat. Genet. 35:32-40(2003).
DR	EMBL; BX640449; CAE34379.1; -.
KW	Cell division; Complete proteome.
SQ	SEQUENCE 83 AA; 9489 MW; 98431A7ABDD2F58 CRC64;
QY	Query Match 100.0%; Score 34; DB 16; Length 83;
	Best Local Similarity 100.0%; Pred. No. 16;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DG	
DB	31 GRGDSP 36
RESULT 3	
Q7M4S7	PREIMINARY; PRT; 83 AA.
ID	AC Q7M4S7
AD	DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT	DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT	DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Putative cell division topological specificity factor.
GN	MINE OR BP33581.
OS	Bordetella parapertussis.
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC	Alcaligenaceae; Bordetella.
CX	NCBI_TaxID=519;
RN	[...]
RP	SEQUENCE FROM N.A.
RC	STRAIN=12822 / ATCC BAA-587;
RX	MEDLINE=22827954; PubMed=12910271;
RA	Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA	Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA	Leather S., Moule S., Norberczak E., O'Neill S., Ormond D., Price C.,
RA	Rabinowitz E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,
RT	Bordetella parapertussis and Bordetella bronchiseptica.";
RL	Nat. Genet. 35:32-40(2003).
DR	EMBL; BX640434; CAE38865.1; -.
KW	Cell division; Complete proteome.
SQ	SEQUENCE 83 AA; 9489 MW; 98431A7ABDD2F58 CRC64;
QY	Query Match 100.0%; Score 34; DB 16; Length 83;
	Best Local Similarity 100.0%; Pred. No. 16;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DG	
DB	31 GRGDSP 36

```

QY      1 GRGDSP 6
DB      35 GRGDSP 40

RESULT 6
Q92K28 ID Q92K28 PRELIMINARY; PRT; 208 AA.
AC Q92K28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein R02299.
GN R02299 OR SMC01549.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portecelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591790; CAC46878.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 208 AA; 21420 MW; A9222A33391FB522 CRC64;

Query Match 100.0%; Score 34; DB 16; Length 208;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDSP 6
DB      122 GRGDSP 127

RESULT 7
Q9XSG0 ID Q9XSG0 PRELIMINARY; PRT; 293 AA.
AC Q9XSG0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibronectin (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic smooth muscle;
RA Reichenberg S., Plenz G., Breithardt G.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135404; AAD24483.1; -.
DR HSPF; P02751; iFNH.
DR GO; GO:0016491; iFNH.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SMC0060; FN3; 3.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Repeat.

```

```

FT NON_TER 1
FT NON_TER 293
SQ SEQUENCE 293 AA; 31537 MW; FF0252D9AE853061 CRC64;

Query Match 100.0%; Score 34; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDSP 6
DB      62 GRGDSP 67

RESULT 8
Q8CCB5 ID Q8CCB5 PRELIMINARY; PRT; 363 AA.
AC Q8CCB5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK033488; BAC28315.1; -.
KW Hypothetical protein.
SQ SEQUENCE 363 AA; 40631 MW; 0FDFED6C04E9DB38 CRC64;

Query Match 100.0%; Score 34; DB 11; Length 363;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDSP 6
DB      42 GRGDSP 47

RESULT 9
Q9HQN2 ID Q9HQN2 PRELIMINARY; PRT; 441 AA.
AC Q9HQN2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Isochorismate synthase.
GN MENF OR VNG1083G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11381).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504493; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Keller D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Kellar K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

```

```

DR EMBL; AE005040; AAC9481.1; -.
DR PIR; E84264; E84264.
DR GO; GO:0008909; F:isochorismate synthase activity; IEA.
DR GO; GO:0069058; P:biosynthesis; IEA.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR004561; Isochor_synth.
DR Pfam; PF02425; chorismate_bind; 1.
DR ProDom; PD000779; Anth synth chor; 1.
DR TIGRFAMs; TIGR00543; isochor_syn; 1.
KW Complete proteome.
SQ SEQUENCE 441 AA; 46011 MW; A98F51217A93B98C CRC64;

Query Match 100.0%; Score 34; DB 17; Length 441;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSF 6
Db 271 GRGDSF 276

RESULT 10
QBNC49 PRELIMINARY; PRT; 562 AA.
AC QBNC49;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein FLJ30494.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Negai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yanamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Minomiya K.;
RC "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074975; BAC11326.1; -.
KW Hypothetical protein.
SQ SEQUENCE 562 AA; 61286 MW; 708041459E34D78A CRC64;

Query Match 100.0%; Score 34; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSF 6
Db 486 GRGDSF 491

RESULT 11
QBTVQ5 PRELIMINARY; PRT; 618 AA.
AC QBTVQ5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Unclassified secreted protein.
GN MK237.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AV19 / DSM 6324 / JCM 9639;
RC MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,

```

```

RA Stcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Nyalale D.A., Rogozin I.B., Tatuzov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlyak S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
EL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010321; AM01454.1; -.
KW Complete proteome.
SQ SEQUENCE 618 AA; 68154 MW; 168BA3872D50BBEB CRC64;

Query Match 100.0%; Score 34; DB 17; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSF 6
Db 553 GRGDSF 558

RESULT 12
QBGMX7 PRELIMINARY; PRT; 659 AA.
AC QBGMX7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein (At3g53930).
GN AT3G53930/F5K20.230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlic-Reumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Sarou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK119572; BAC43172.1; -.
DR EMBL; BT005945; AAO64880.1; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0016831; P:carboxy-lyase activity; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.

```

DR PROSITE; PSC0392; DCC GAD HDC YDC; 1.  
 DR PROSITE; PSC0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSC0108; PROTEIN\_KINASE\_ST; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 659 AA; 73464 MW; 00AA484E02D62C34 CRC64;  
 Query Match 100.0%; Score 34; DB 10; Length 659;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 DB 357 GRGDSP 362  
 RESULT 13  
 Q9M334 PRELIMINARY; PRT; 691 AA.  
 AC Q9M334;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE Hypothetical protein.  
 GN F5K20.230.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RA Monfort A., Casacuberta E., Puigdomenech P., Mewes H.W., Rudd S.,  
 RA Lemcke K., Mayer K.F.X., Quetier F., Salasoubat M.;  
 RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132960; CAB80355.1; -;  
 DR PIR; T45933; T45933.  
 DR HSP; O63450; 1A06.  
 DR GO; GO:000524; F:ATP binding; IEA.  
 DR GO; GO:0016831; F:carboxy-lyase activity; IEA.  
 DR GO; GO:0004672; F:protein kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR GO; GO:0015992; P:proton transport; IEA.  
 DR InterPro; IPR000194; ATPase\_a/bcentre.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002129; Pyridoxal\_dec.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; 1.  
 DR PROSITE; PSC0392; DCC GAD HDC YDC; 1.  
 DR PROSITE; PSC00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PSC0011; PROTEIN\_KINASE\_DOM; 1.  
 KW Hypothetical protein; ATP-binding; Transferase.  
 SQ SEQUENCE 691 AA; 76711 MW; D04A2660DA1393EB CRC64;  
 Query Match 100.0%; Score 34; DB 10; Length 691;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 DB 389 GRGDSP 394  
 RESULT 14  
 O93406 PRELIMINARY; PRT; 2478 AA.  
 ID O93406  
 AC O93406;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Fibronectin.  
 GN FNI.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RA Zhao Q., Colloidi P.;  
 RA "Characterization and expression of zebrafish fibronectin."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF081128; AAC31947.1; -;  
 DR HSP; P02751; IFBR  
 DR ZFIN; ZDB-GENE-000426-1; fnl;  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR InterPro; IPR009562; FN\_Type\_II.  
 DR Pfam; PF00039; fnl; 12.  
 DR Pfam; PF00040; fn2; 2.  
 DR Pfam; PF00041; fn3; 17.  
 DR PRINTS; PR00013; FNTYPRII.  
 DR ProDom; PD000995; FN\_Type\_II; 2.  
 DR SMART; SM00058; FN1; 2.  
 DR SMART; SM00059; FN2; 2.  
 DR SMART; SM00060; FN3; 14.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 11.  
 DR PROSITE; PS03023; FIBRONECTIN\_2; 2.  
 SQ SEQUENCE 2478 AA; 271652 MW; A03475C5A385750 CRC64;  
 Query Match 100.0%; Score 34; DB 13; Length 2478;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSP 6  
 DB 1613 GRGDSP 1618  
 RESULT 15  
 Q882M6 PRELIMINARY; PRT; 3173 AA.  
 AC Q882M6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Yersiniabactin polyketide/non-ribosomal peptide synthetase.  
 GN YRP1 OR PSPT02600.  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=PC3000;  
 RA Buell R., Ueerdback V., Khouri H., Fedorova N., Tran B., Russell D.,  
 RA Berry K., Ueerdback T., Van Aken S., Feldblyum T., Gwinn M.,  
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,  
 RA White O., Fraser C., Collier A.;  
 RA "Complete sequence of Pseudomonas syringae."  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE016865; AA056104.1; -;  
 DR TIGR; PSPT02600; -;  
 DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.

Fri Apr 16 11:04:17 2004

DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR InterPro; IPR001227; AC trans.  
 DR InterPro; IPR001242; Cordensatn.  
 DR InterPro; IPR000794; Ketoacyl synth.  
 DR InterPro; IPR001601; Methyl-transf.  
 DR InterPro; IPR006162; Ppantne S.  
 DR InterPro; IPR006163; Pp bind.  
 DR InterPro; IPR000051; SAM bind.  
 DR InterPro; IPR001031; Thioesterase.  
 DR Pfam; PF00698; Acyl transf; 1.  
 DR Pfam; PF00668; Condensation; 1.  
 DR Pfam; PF00109; ketoacyl-synt; 1.  
 DR Pfam; PF02801; ketoacyl-synt C; 1.  
 DR Pfam; PF00550; pp-binding; 2.  
 DR Pfam; PF00975; Thioesterase; 1.  
 DR PROSITE; PS00075; ACP DOMAIN; 2.  
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.  
 KW Complete proteome.  
 SQ SEQUENCE 3173 AA; 348230 MW; 4B87D3E2B747F1FE CRC64;

Query Match 100.0%; Score 34; DB 16; Length 3173;  
 Best Local Similarity 100.0%; Pred. No. 7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6  
 |||||  
 Db 2222 GRGDSP 2227

Search completed: April 16, 2004, 07:27:49  
 Job time : 49.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:28:08 ; Search time 40.35 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588B-8

Perfect score: 35

Sequence: 3 GRGDTF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/ECTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09D\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	6	11 US-09-991-588B-8	Sequence 8, Appli
2	35	100.0	6	12 US-09-780-612A-1	Sequence 1, Appli
3	35	100.0	6	12 US-09-961-381A-4	Sequence 4, Appli
4	35	100.0	6	13 US-10-007-270-33	Sequence 33, Appli
5	35	100.0	6	14 US-10-263-456-1	Sequence 1, Appli
6	35	100.0	6	14 US-10-420-029-1	Sequence 1, Appli
7	35	100.0	344	12 US-10-425-114-63361	Sequence 63361, A
8	35	100.0	354	12 US-10-425-114-63236	Sequence 63236, A
9	35	100.0	367	12 US-10-425-114-62790	Sequence 62790, A
10	35	100.0	379	12 US-10-425-114-60653	Sequence 60653, A
11	35	100.0	477	12 US-10-425-114-71233	Sequence 71233, A
12	35	100.0	1005	14 US-10-193-764-41	Sequence 41, Appli
13	35	100.0	1011	14 US-10-193-764-39	Sequence 39, Appli
14	32	91.4	1226	15 US-10-369-493-5807	Sequence 5807, Ap
15	32	91.4	1517	14 US-10-156-761-13265	Sequence 13265, A

16	31	88.6	6	9 US-09-837-969A-28	Sequence 28, Appli
17	31	88.6	6	9 US-09-837-969A-46	Sequence 46, Appli
18	31	88.6	6	9 US-09-888-260-15	Sequence 15, Appli
19	31	88.6	6	9 US-09-961-834-1	Sequence 1, Appli
20	31	88.6	6	9 US-09-364-597A-21	Sequence 21, Appli
21	31	88.6	6	9 US-09-177-847-1	Sequence 1, Appli
22	31	88.6	6	9 US-09-841-321A-28	Sequence 28, Appli
23	31	88.6	6	9 US-09-841-321A-46	Sequence 46, Appli
24	31	88.6	6	9 US-09-871-974-3	Sequence 3, Appli
25	31	88.6	6	9 US-09-972-772-31	Sequence 31, Appli
26	31	88.6	6	11 US-09-991-588B-7	Sequence 7, Appli
27	31	88.6	6	12 US-09-961-381A-3	Sequence 3, Appli
28	31	88.6	6	12 US-09-380-682-63	Sequence 63, Appli
29	31	88.6	6	12 US-10-136-187-42	Sequence 42, Appli
30	31	88.6	6	13 US-10-003-945-31	Sequence 31, Appli
31	31	88.6	6	13 US-10-007-270-32	Sequence 32, Appli
32	31	88.6	6	13 US-10-007-270-35	Sequence 35, Appli
33	31	88.6	6	13 US-10-114-176-2	Sequence 2, Appli
34	31	88.6	6	14 US-10-046-901-5	Sequence 5, Appli
35	31	88.6	6	14 US-10-237-850-83	Sequence 83, Appli
36	31	88.6	6	14 US-10-138-935-31	Sequence 31, Appli
37	31	88.6	6	14 US-10-262-435-5	Sequence 5, Appli
38	31	88.6	6	14 US-10-325-021-9	Sequence 9, Appli
39	31	88.6	6	14 US-10-323-013-2	Sequence 2, Appli
40	31	88.6	6	14 US-10-439-532-15	Sequence 15, Appli
41	31	88.6	6	14 US-10-086-208-5	Sequence 5, Appli
42	31	88.6	6	15 US-10-259-609-14	Sequence 14, Appli
43	31	88.6	6	15 US-10-182-432-31	Sequence 31, Appli
44	31	88.6	7	8 US-08-987-756-1	Sequence 1, Appli
45	31	88.6	7	9 US-09-364-597A-22	Sequence 22, Appli

## ALIGNMENTS

### RESULT 1

US-09-991-588B-8

Sequence 8, Application US/09991588B

Publication No. US20030219429A1

GENERAL INFORMATION:

APPLICANT: Budny, John A.

TITLE OF INVENTION: Compositionand Method for Bone Regeneration

FILE REFERENCE: 1008-130.US

CURRENT APPLICATION NUMBER: US/09/991,588B

CURRENT FILING DATE: 2001-11-21

PRIOR APPLICATION NUMBER: US 09/122,348

PRIOR FILING DATE: 1998-07-24

NUMBER OF SEQ ID NOS: 23

SEQ ID NO 8

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Purchased commercially or sequence is synthesized

US-09-991-588B-8

Query Match 100.0%; Score 35; DB 11; Length 6;  
Best Local Similarity 100.0%; Pred. No. 16+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTF 6

Db 1 GRGDTF 6

### RESULT 2

US-09-780-612A-1

Sequence 1, Application US/09780612A

Publication No. US20010053766A1

GENERAL INFORMATION:

APPLICANT: KUMAR, JANARDAN

APPLICANT: RAO, VASANTH

APPLICANT: EPSTEIN, DAVID S.

;; TITLE OF INVENTION: METHOD OF TREATING DISORDERS OF THE EYE

;; FILE REFERENCE: 1579-434

;; CURRENT APPLICATION NUMBER: US/09/780,612A

;; CURRENT FILING DATE: 2001-02-12

;; PRIOR APPLICATION NUMBER: 50/181,869

;; PRIOR FILING DATE: 2000-02-11

;; NUMBER OF SEQ ID NOS: 2

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 1

;; TYPE: PRT

;; ORGANISM: Artificial Sequence

;; FEATURE:

;; OTHER INFORMATION: Description of Artificial Sequence

US-09-780-612A-1

Query Match 100.0%; Score 35; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 1e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTG 6

Db 1 GRGDTG 6

RESULT 3

US-09-961-381A-4

;; Sequence 4, Application US/09961381A

;; Publication No. US20020061515A1

;; GENERAL INFORMATION:

;; APPLICANT: Lynch, Gary

;; APPLICANT: Bi, Xiaoning

;; APPLICANT: Gall, Christine M.

;; TITLE OF INVENTION: Model for Neurodegenerative Diseases Involving Amyloid

;; FILE REFERENCE: 1819.0040001

;; CURRENT APPLICATION NUMBER: US/09/961,381A

;; CURRENT FILING DATE: 2001-09-25

;; PRIOR APPLICATION NUMBER: US 60/235,374

;; PRIOR FILING DATE: 2000-09-25

;; NUMBER OF SEQ ID NOS: 5

;; SOFTWARE: PatentIn version 3.2

;; SEQ ID NO 4

;; LENGTH: 6

;; TYPE: PRT

;; ORGANISM: Unknown

;; FEATURE:

;; OTHER INFORMATION: Synthetic Peptide

US-09-961-381A-4

Query Match 100.0%; Score 35; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 1e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTG 6

Db 1 GRGDTG 6

RESULT 4

US-10-007-270-33

;; Sequence 33, Application US/10007270

;; Publication No. US20020160954A1

;; GENERAL INFORMATION:

;; APPLICANT: Hageman, Gregory S.

;; APPLICANT: Kaehn, Markus H.

;; APPLICANT: University of Iowa Research Foundation

;; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR OCULAR ABNORMALITIES

;; FILE REFERENCE: 020618-000120US

;; CURRENT APPLICATION NUMBER: US/10/007,270

;; CURRENT FILING DATE: 2001-11-08

;; PRIOR APPLICATION NUMBER: US 09/430,195

;; PRIOR FILING DATE: 1999-10-29

;; PRIOR APPLICATION NUMBER: US 09/183,972

;; PRIOR FILING DATE: 1998-10-29

;; NUMBER OF SEQ ID NOS: 37

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 33

;; LENGTH: 6

;; TYPE: PRT

;; ORGANISM: Artificial Sequence

;; FEATURE:

;; OTHER INFORMATION: Description of Artificial Sequence

;; OTHER INFORMATION: RGD-containing peptide

US-10-007-270-33

Query Match 100.0%; Score 35; DB 13; Length 6;

Best Local Similarity 100.0%; Pred. No. 1e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTG 6

Db 1 GRGDTG 6

RESULT 5

US-10-263-456-1

;; Sequence 1, Application US/10263456

;; Publication No. US20030077646A1

;; GENERAL INFORMATION:

;; APPLICANT: GREG BIESECKER

;; SUMEDHA D. JAYASENA

;; LARRY GOLD

;; DREW SMITH

;; GARY P. KIRSCHENHEUTER

;; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY

EXPONENTIAL ENRICHMENT: BLENDED

SELEX

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Swanson & Bratschun, L.L.C.

STREET: 1745 Shea Center Drive, Suite 330

CITY: Highlands Ranch

STATE: Colorado

COUNTRY: USA

ZIP: 80129

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage

COMPUTER: IBM pc compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/263,456

FILING DATE: 02-Oct-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/606,477

FILING DATE: 29-Jun-2000

APPLICATION NUMBER: 08/856,699

FILING DATE: OCTOBER 23, 1997

APPLICATION NUMBER: 08/234,997

FILING DATE: APRIL 28, 1994

APPLICATION NUMBER: 07/714,131

FILING DATE: JUNE 10, 1991

APPLICATION NUMBER: 07/536,428

FILING DATE: JUNE 11, 1990

APPLICATION NUMBER: 08/117,991

FILING DATE: SEPTEMBER 8, 1993

APPLICATION NUMBER: 08/123,935

FILING DATE: SEPTEMBER 17, 1993

APPLICATION NUMBER: 08/199,507

FILING DATE: FEBRUARY 22, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Barry J. Swanson

REGISTRATION NUMBER: 33,215

REFERENCE/DOCKET NUMBER: NEX15/C-CON

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-263-456-1

Query Match 100.0%; Score 35; DB 14; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
|||||  
Db 1 GRGDTP 6

## RESULT 6

US-10-420-029-1  
Sequence 1, Application US/10420029  
Publication No. US20030157712A1  
GENERAL INFORMATION:  
APPLICANT: Daniel, Thomas O.  
TITLE OF INVENTION: Methods for Determining Cell Responses  
FILE REFERENCE: 22000.008502  
CURRENT APPLICATION NUMBER: US/10/420,029  
CURRENT FILING DATE: 2003-04-17  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: 09/485,653  
PRIOR FILING DATE: 1998-08-19  
PRIOR APPLICATION NUMBER: PCT/US98/17157  
PRIOR FILING DATE: 1998-08-19  
PRIOR APPLICATION NUMBER: 60/056,164  
PRIOR FILING DATE: 1997-08-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence; No. US20030157712A1e =  
US-10-420-029-1

Query Match 100.0%; Score 35; DB 14; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
|||||  
Db 1 GRGDTP 6

## RESULT 7

US-10-425-114-63361  
Sequence 63361, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 63361  
LENGTH: 344  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLM017220E10\_FLI.pep  
US-10-425-114-63361

Query Match 100.0%; Score 35; DB 12; Length 344;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
|||||  
Db 96 GRGDTP 101

## RESULT 8

US-10-425-114-63236  
Sequence 63236, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 63236  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3180-021-H3\_FLI.pep  
US-10-425-114-63236

Query Match 100.0%; Score 35; DB 12; Length 354;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
|||||  
Db 103 GRGDTP 108

## RESULT 9

US-10-425-114-62790  
Sequence 62790, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 62790  
LENGTH: 367



```
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3355-012-E1_FLI.pep
US-10-425-114-62790

Query Match      100.0%; Score 35; DB 12; Length 367;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRGDTp 6
Db      117 GRGDTp 122

RESULT 1C
US-10-425-114-60653
; Sequence 60653, Application US/10425114
; Publication No. US200340034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60653
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3357-014-C4_FLI.pep
US-10-425-114-60653

Query Match      100.0%; Score 35; DB 12; Length 379;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRGDTp 6
Db      103 GRGDTp 108

RESULT 11
US-10-425-114-71233
; Sequence 71233, Application US/10425114
; Publication No. US200340034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71233
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2XFLM017004F07_FLI.pep
US-10-425-114-71233
```

```
Query Match      100.0%; Score 35; DB 12; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRGDTp 6
Db      107 GRGDTp 112

RESULT 12
US-10-193-764-41
; Sequence 41, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1338-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-41

Query Match      100.0%; Score 35; DB 14; Length 1005;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRGDTp 6
Db      19 GRGDTp 24

RESULT 13
US-10-193-764-39
; Sequence 39, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1338-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-39

Query Match      100.0%; Score 35; DB 14; Length 1011;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GRGDTp 6
Db      25 GRGDTp 30
```

RESULT 14  
 US-10-369-493-5807  
 ; Sequence 5807, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 5807  
 ; LENGTH: 1226  
 ; TYPE: PRT  
 ; ORGANISM: Caenorhabditis elegans  
 US-10-369-493-5807

Query Match 91.4%; Score 32; DB 15; Length 1226;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |:|||||  
 Db 519 GKGDTP 524

RESULT 15  
 US-10-156-761-13265  
 ; Sequence 13265, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15139  
 ; SEQ ID NO 13265  
 ; LENGTH: 1517  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 US-10-156-761-13265

Query Match 91.4%; Score 32; DB 14; Length 1517;  
 Best Local Similarity 83.3%; Pred. No. 1.5e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |:|||||  
 Db 512 GKGDTP 1517

Search completed: April 16, 2004, 08:03:16  
 Job time : 40.35 secs

GenCore version: 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time: 12.6 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-8

Perfect score: 35

Sequence: 1 GRGDTF 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 3%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	35	100.0	370	2 S27344	hupK protein - Rhizobium leguminosarum
2	35	100.0	450	2 D87342	multidrug resistance protein, probable [imported] - Caulobacter crescentus
3	35	100.0	1307	2 T10887	-46D nuclear prote
4	32	91.4	266	1 S18159	ribosomal protein
5	32	91.4	458	2 T02571	probable myrosinase
6	32	91.4	606	2 AC2508	hypothetical prote
7	32	91.4	1226	2 S44824	FS4F2.1 protein -
8	31	88.6	273	2 A28512	fibronectin - chic
9	31	88.6	399	2 G84270	hypothetical prote
10	31	88.6	425	1 S48469	probable membrane
11	31	88.6	438	2 B82042	osmolality sensor
12	31	88.6	441	2 E84264	isochromatase synt
13	31	88.6	463	1 S74845	tldd homolog sir08
14	31	88.6	515	2 A95675	probable oxidoredu
15	31	88.6	515	2 D85525	probable oxidoredu
16	31	88.6	606	2 T29190	hypothetical prote
17	31	88.6	691	2 T45933	hypothetical prote
18	31	88.6	1020	2 A29355	fibronectin - chic
19	31	88.6	2265	1 FN80	fibronectin - bovi
20	31	88.6	2386	1 FNHU	fibronectin precu
21	31	88.6	2477	2 S14428	fibronectin - Afri
22	31	88.6	2481	2 A43908	fibronectin - Afri
23	30	85.7	253	2 F84258	hypothetical prote
24	30	85.7	277	2 A03256	aminoacylase (EC 3
25	30	85.7	299	2 C55213	rfbC protein - Shi
26	30	85.7	317	2 A75477	cysteine synthase
27	30	85.7	335	1 KHR0B	cathepsin B (EC 3.
28	30	85.7	339	1 KHRUB	cathepsin B (EC 3.
29	30	85.7	339	1 KHRTB	cathepsin B (EC 3.

## ALIGNMENTS

### RESULT 1

S27344

hupK protein - Rhizobium leguminosarum

C:Species: Rhizobium leguminosarum

C>Date: 25-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 22-Jun-2003

C:Accession: S27344

R:Ref: L.; Hidalgo, E.; Palacios, J.; Ruiz-Argueso, T.

J. Mol. Biol. 228, 998-1002, 1992

A:Title: Nucleotide sequence and organization of an H(2)-uptake gene cluster from Rhizob

A:Reference number: S27340; MUID:93108466; PMID:1469733

A:Accession: S27344

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-370 <REY>

A:Cross-references: EMBL:X52974; NID:gl167855; PIDN:CAA37158.1; PID:g48731

C:Superfamily: hydrogenase maturation factor, hupK type

Query Match 100.0%; Score 35; DB 2; Length 370;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTF 6

Db 163 GRGDTF 168

### RESULT 2

D87342

multidrug resistance protein, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: D87342

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapero, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: D87342

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-450 <STO>

A:Cross-references: GB:AE005673; NID:gl3421986; PIDN:AAK22736.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0751

Query Match 100.0%; Score 35; DB 2; Length 450;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTF 6

|||||

Db 220 GRGDTP 225

## RESULT 3

T30887  
145D nuclear protein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T30887  
R:Schmidt-Zachmann, M.S.; Knecht, S.; Kraemer, A.  
Mol. Biol. Cell 9, 143-160, 1998  
A:Title: Molecular characterization of a novel, widespread nuclear protein that co-localizes with the nucleolus  
A:Reference number: Z2832; MUID:98099695; PMID:9436997  
A:Accession: T30887  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1307 <SCK>  
A:Cross-references: EMBL:Y08997; NID:G2791895; PIDN:CAA70201.1; PID:G2791896

Query Match 100.0%; Score 35; DB 2; Length 1307;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 272 GRGDTP 277

## RESULT 4

S18159  
ribosomal protein L7a, cytosolic - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 22-Jun-1999  
C:Accession: A46032; I50416; S18159  
R:Colombo, P.; Yon, J.; Garson, K.; Fried, M.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6358-6362, 1992  
A:Title: Conservation of the organization of five tightly clustered genes over 600 million years  
A:Reference number: A46032; MUID:92335297; PMID:1631131  
A:Accession: A46032  
A:Molecule type: mRNA; DNA  
A:Residues: 1-266 <COL>  
A:Cross-references: EMBL:X62640; NID:G63777; PIDN:CAA44506.1; PID:G63778  
A:Note: submitted to the EMBL Data Library, October 1991  
A:Note: sequence extracted from NCBI backbone (NCBI:108733)

R:Maeda, N.; Kenmochi, N.; Tanaka, T.  
Biochimie 75, 785-790, 1993  
A:Title: The complete nucleotide sequence of chicken ribosomal protein L7a gene and the corresponding cDNA  
A:Reference number: I50416; MUID:94100270; PMID:8274530  
A:Accession: I50416  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-266 <MAE>  
A:Cross-references: GB:D14522; NID:9457652; PIDN:BA03395.1; PID:G601884

C:Genetics:  
A:Gene: surf-3/Rpl7a  
A:Introns: 1/3; 42/1; 92/1; 139/1; 165/3; 209/2; 232/3  
C:Superfamily: rat ribosomal protein L7a  
C:Keywords: protein biosynthesis; ribosome

Query Match 91.4%; Score 32; DB 1; Length 266;  
Best Local Similarity 83.3%; Pred. No. 32;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 124 GRGDTP 129

## RESULT 5

T02571  
probable myosinase-binding protein [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein T16B24.5  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
C:Accession: T02571; G84815  
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, August 1998  
A:Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.

A:Reference number: Z14679

A:Accession: T02571

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-458 &lt;ROU&gt;

A:Cross-references: EMBL:AC004697; NID:G3402671; PID:G3402676

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Sten, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84815

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-458 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:G3402676; PIDN:AAC28979.1; GSPDB:GN00139

C:Genetics:

A:Gene: T16B24.5; A12G39310

A:Map position: 2

A:Introns: 67/3; 221/3; 374/3

Query Match 91.4%; Score 32; DB 2; Length 458;  
Best Local Similarity 83.3%; Pred. No. 57;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 214 GRGDTP 219

## RESULT 6

AC2508  
hypothetical protein alr7243 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AC2508  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AC2508  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-606 <KUR>  
A:Cross-references: GB:BA000020; PIDN:BA878327.1; PID:G17135781; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr7243

A:Genome: plasmid

Query Match 91.4%; Score 32; DB 2; Length 606;  
Best Local Similarity 83.3%; Pred. No. 76;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 120 GRGDTP 125

## RESULT 7

S44824  
F54F2.1 protein - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 14-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 23-Mar-2001

C;Accession: S44824  
 R;Anderson, K.  
 submitted to the EMBL Data Library, September 1993  
 A;Description: Sequence of the C. elegans cosmid F54F2.  
 A;Reference number: S44817  
 A;Accession: S44824  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1226 <AND>  
 A;Cross-references: EMBL:L23645; NID:g338603; PID:g388605  
 C;Genetics:  
 A;Introns: 58/2; 137/3; 179/1; 316/2; 393/1; 551/3; 597/2; 662/2; 899/3; 1178/3  
 C;Keywords: cytoskeleton; transmembrane protein

Query Match 91.4%; Score 32; DB 2; Length 1226;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |||||  
 DB 519 GRGDTP 524

RESULT 8  
 A28512  
 fibronectin - chicken (fragment)  
 C;Species: Gallus gallus (chicken)  
 C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Aug-1999  
 C;Accession: A28512  
 R;Kubocura, S.; Obara, M.; Karasaki, Y.; Taniguchi, H.; Gotoh, S.; Tsuda, T.; Higashi, K.  
 Biochim. Biophys. Acta 910, 171-181, 1987  
 A;Title: Genetic analysis of the cell binding domain region of the chicken fibronectin  
 A;Reference number: A28512; MUID:88050950; PMID:2823899  
 A;Accession: A28512  
 A;Molecule type: DNA  
 A;Residues: 1-273 <KUB>  
 A;Cross-references: GB:X04533; NID:g63393; PIDN:CAA29781.1; PID:g295716  
 A;Note: the authors translated the codon CCG for residue 190 as Gln, CAG for residue 243  
 C;Genetics:  
 A;Introns: 90/1; 129/1; 184/1; 236/1  
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat  
 C;Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; heter  
 F;1-82/Domain: fibronectin type III repeat homology (fragment) <FN3I>  
 F;90-172/Domain: fibronectin type III repeat homology <FN3J>  
 F;167-169/Region: cell attachment (R-G-D) motif  
 F;184-266/Domain: fibronectin type III repeat homo.cgy <FN3K>

Query Match 88.6%; Score 31; DB 2; Length 273;  
 Best Local Similarity 83.3%; Pred. No. 53;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |||||  
 DB 166 GRGDSP 171

RESULT 9  
 G84270  
 hypothetical protein Vngl149c [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C;Accession: G84270  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 J. Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; L  
 A;Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: G84270  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-399 <STO>

A;Cross-references: GB:AE004437; NID:gl0580687; PIDN:AAAG19531.1; GSPDB:GN00138  
 C;Genetics:  
 A;Gene: VNG1149C

Query Match 88.6%; Score 31; DB 2; Length 399;  
 Best Local Similarity 83.3%; Pred. No. 80;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |||||  
 DB 287 GRGETP 292

RESULT 10  
 S48469  
 probable membrane protein YIL103w - yeast (Saccharomyces cerevisiae)  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Apr-2002  
 C;Accession: S48469  
 R;Bowman, S.; Churcher, C.  
 submitted to the EMBL Data Library, September 1994  
 A;Reference number: S48455  
 A;Accession: S48469  
 A;Molecule type: DNA  
 A;Residues: 1-425 <BOW>  
 A;Cross-references: GB:Z47047; EMBL:Z38125; NID:g633997; PID:g763243; GSPDB:GN00009; MI  
 C;Genetics:  
 A;Gene: MIP5; YIL103w  
 A;Cross-references: SGD:S0001365  
 A;Map position: 9L  
 C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1803  
 C;Keywords: transmembrane protein  
 F;155-171/Domain: transmembrane #status predicted <TM1>  
 F;334-350/Domain: transmembrane #status predicted <TM2>

Query Match 88.6%; Score 31; DB 1; Length 425;  
 Best Local Similarity 83.3%; Pred. No. 85;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |||||  
 DB 415 GRGETP 420

RESULT 11  
 B82042  
 osmolarity sensor protein EnvZ VC2713 [imported] - Vibrio cholerae (strain N16961 serog  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: B82042  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J  
 I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.W.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: B82042  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-438 <HEI>  
 A;Cross-references: GB:AE004337; GB:AE003852; NID:g9657307; PIDN:AAF95853.1; GSPDB:GN001  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC2713  
 A;Map position: 1

Query Match 88.6%; Score 31; DB 2; Length 438;  
 Best Local Similarity 83.3%; Pred. No. 88;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 |||||  
 DB 193 GRGETP 195

```

RESULT 12
E84264
isochorismate synthase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 12-Jun-2003
C/Accession: E84264
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, E.D.; Lasky, S.;
; Leitthausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: E84264
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-441 <STO>
A/Cross-references: GB:AE004437; NID:gl0580629; PIDN:RAG:9481.1; GSPDB:GN00138
C/Genetics:
A/Gene: menF
C:Superfamily: isochorismate synthase

Query Match      88.6%; Score 31; DB 2; Length 441;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      271 GRGDSP 276

RESULT 13
S74845
tldd homolog slr0863 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A/Variety: PCC 6803
C>Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 16-Jun-2000
C/Accession: S74845
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shino, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 129-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S74845
A/Molecule type: DNA
A/Residues: 1-463 <KAN>
A/Cross-references: EMBL:D90909; GB:AB001339; NID:gl652844; PIDN:BAAL7806.1; PID:gl65284
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: Escherichia coli tldd protein

Query Match      88.6%; Score 31; DB 1; Length 463;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      352 GRGETP 357

RESULT 14
A99675
probable oxidoreductase subunit Ecs0369 [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: A99675
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shitagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796

```

```

A/Accession: A99675
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-515 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA833792.1; PID:gl3359826; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: Ecs0369

Query Match      88.6%; Score 31; DB 2; Length 515;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      284 GRGETP 289

RESULT 15
D85525
probable oxidoreductase subunit yahF [imported] - Escherichia coli (strain O157:H7, sub
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: D85525
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotback, E.J.; Davis, N.W.; Lim, A.; Dimallanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: D85525
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-515 <STO>
A/Cross-references: GB:AE005174; NID:gl2513137; PIDN:RAG54664.1; GSPDB:GN00145; UWGP:Z0
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: yahF

Query Match      88.6%; Score 31; DB 2; Length 515;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTP 6
      |||||
Db      284 GRGETP 289

Search completed: April 16, 2004, 07:29:24
Job time : 14.6 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: April 16, 2004, 06:56:42 ; Search time 10.5 Seconds  
(without alignments)  
29.754 Millior cell updates/sec

Title: US-09-991-588E-8

Perfect score: 35

Sequence: 1 GRGDTF 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Preð. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	370	1 HUPK RHILV	P28153 rhizobium l
2	35	100.0	1304	1 S3BI HUMAN	O75533 homo sapien
3	35	100.0	1304	1 S3BI MOUSE	O99nb9 mus musculus
4	35	100.0	1307	1 S3BI XENLA	O57683 xenopus lae
5	32	91.4	265	1 RLVA CHICK	P32429 gallus galli
6	32	91.4	458	1 MS22 ARATH	O80950 arabidopsis
7	32	91.4	1226	1 PAT2 CAEE-	P34446 caenorhabdi
8	31	88.6	425	1 YIK3 YAST	P40487 saccharomyc
9	31	88.6	463	1 Y863 SYN3	P73754 synechocyst
10	31	88.6	1256	1 FINC CHICK	P11722 gallus galli
11	31	88.6	1328	1 FINC FLEWA	O91289 pleurodeles
12	31	88.6	2265	1 FINC BOVIN	P07589 bos taurus
13	31	88.6	2386	1 FINC HUMAN	P02751 homo sapien
14	31	88.6	2477	1 FINC MOUSE	P11276 mus musculus
15	31	88.6	2477	1 FINC RAT	P04937 rattus norv
16	31	88.6	2481	1 FINC XENLA	O91740 xenopus lae
17	30	85.7	253	1 GC32 HALNI	O9hgs9 haibacteri
18	30	85.7	299	1 RFBD SHIFL	P37778 sh-gellia fl
19	30	85.7	335	1 CATE BOVIN	P07688 bos taurus
20	30	85.7	339	1 CATE HUMAN	P07858 homo sapien
21	30	85.7	339	1 CATE MOUSE	P10605 mus musculus
22	30	85.7	339	1 CATE RAT	P00787 rattus norv
23	30	85.7	499	1 ACH3 STRCO	P04757 rattus norv
24	30	85.7	580	1 NADB TRACO	O9xh8 streptomyce
25	30	85.7	655	1 HS70 DAUCA	P26791 daucus caro
26	30	85.7	746	1 RED2 RAT	P97616 rattus norv
27	30	85.7	1030	1 SM6A HUMAN	O9h2e6 homo sapien
28	29	82.9	61	1 CUB9 HUMAN	O96h27 homo sapien
29	29	82.9	160	1 R1SB MYCLE	O9ccp3 mycobacteri
30	29	82.9	216	1 RSEA ECOLI	P38106 escherichia
31	29	82.9	235	1 NHAB RHOSO	O53117 rhodococcus
32	29	82.9	241	1 CRTA RHCCA	P17055 rhodobacter
33	29	82.9	286	1 RT28 YEAST	P21771 saccharomyc

34 29 82.9 294 1 POL SMSAV  
35 29 82.9 332 1 ACOA ALCEU  
36 29 82.9 343 1 GASI\_MOUSE  
37 29 82.9 345 1 GASI\_HUMAN  
38 29 82.9 387 1 RECF CAUCR  
39 29 82.9 425 1 POU1 BRARE  
40 29 82.9 432 1 LMB1 VIBPA  
41 29 82.9 436 1 TBX6\_HUMAN  
42 29 82.9 443 1 ZP23\_BEARE  
43 29 82.9 467 1 VE2\_HPV24  
44 29 82.9 485 1 ALGI\_PSEPK  
45 29 82.9 493 1 CPE1\_HUMAN

#### ALIGNMENTS

RESULT 1  
HUPK RHILV STANDARD; PRT; 370 AA.  
AC P28153;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Hydrogenase expression/formatation protein hupK.  
GN HUPK.  
OS Rhizobium leguminosarum (biovar viciae).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
CX NCBI\_TaxID=387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=128C53;  
RX MEDLINE=93108466; PubMed=1469733;  
RA Rey L., Hidalgo E., Palacios J.M., Ruiz-Argueso T.;  
RT "Nucleotide sequence and organization of an H2-uptake gene cluster from Rhizobium leguminosarum bv. viciae containing a rubredoxin-like gene and four additional open reading frames.";  
RT J. Mol. Biol. 228:998-1002(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95020662; PubMed=7934943;  
RA Imperial J., Rey L., Ruiz-Argueso T.;  
RT "HupK, a hydrogenase-ancillary protein from Rhizobium leguminosarum, shares structural motifs with the large subunit of Nife hydrogenases and could be a scaffolding protein for hydrogenase metal cofactor assembly.";  
RT Mol. Microbiol. 9:1305-1306(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BL0;  
RA Brito B., Palacios J.M., Imperial J., Ruiz-Argueso T., Yang W.C., Bisseling T., Schmitt H., Kerl V., Bauer T., Kokotek W., Lotz W.;  
RT "Organization of the hup-region and its differential transcription in non-symbiotic and symbiotic cells of Rhizobium leguminosarum bv. viciae BL0.";  
RT Mol. Plant Microbe Interact. 8:235-240(1997).  
RN [4]  
CC -!- SIMILARITY: Belongs to the hupK family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X52974; CAA37158.1; --  
DR EMBL; Z36981; CAA85440.1; --  
DR PIR; S27344; S27344.  
DR PIRSF; PIRSF036413; Hdrgrns\_matr HupK; 1.  
FT CONFLICT 240 240 A -> R (IN REF. 1).  
SQ SEQUENCE 370 AA; 38751 MW; ED17BC76EB2F28C9 CRC64;

```

Query Match      100.0%; Score 35; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTG 6
DB      163 GRGDTG 163

RESULT 2
S3B1_HUMAN
ID      S3B1_HUMAN      STANDARD;      PRT; 1304 AA.
AC      Q75533;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Splicing factor 3B subunit 1 (spliceosome associated protein 155) (SAP
DE      155) (SF3B155) (Pre-mRNA splicing factor SF3B 155 kDa subunit).
GN      SF3B1 OR SAPI55.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RS      SEQUENCE FROM N.A., AND PHOSPHORYLATION.
RX      MEDLINE=98252826; PubMed=9585501;
RA      Wang C., Chua K., Seghezzi W., Lees E., Gorani O., Reed R.;
RT      "Phosphorylation of spliceosomal protein SAP 155 coupled with splicing
RT      catalysis.";
RL      Genes Dev. 12:1409-1414 (1998).
RN      [2]
RS      SEQUENCE OF 1011-1304 FROM N.A.
RE      TISSUE=Brain;
RA      Yu W., Gibbs R.A.;
RN      Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RX      CHARACTERIZATION OF THE SPLICEOSOME.
RA      Das R., Zhou Z., Reed R.;
RT      "Functional association of U2 snRNP with the ATP-independent
RT      spliceosomal complex E.";
RL      Mol. Cell 5:779-787 (2000).
CC      -1- FUNCTION: Subunit of the splicing factor SF3B required for 'A'
CC      complex assembly formed by the stable binding of U2 snRNP to the
CC      branchpoint sequence (bps) in pre-mRNA. Sequence independent
CC      binding of SF3A/SF3B complex upstream of the branch site is
CC      essential, it may anchor U2 snRNP to the pre-mRNA. May also be
CC      involved in the assembly of the 'E' complex. Belongs also to the
CC      minor U2-dependent spliceosome, which is involved in the splicing
CC      of rare class of nuclear pre-mRNA intron.
CC      -1- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF
CC      FOUR SUBUNITS; SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,
CC      SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A
CC      12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS
CC      COMPLEX (U2 snRNP). SF3B1 INTERACTS DIRECTLY WITH THE SPLICING
CC      FACTOR U2AF.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR SPECKLES. DURING MITOSIS,
CC      TRANSIENTLY DISPERSED FROM THE NUCLEAR SPECKLES TO THE CYTOPLASM.
CC      -1- PTM: PHOSPHORYLATED. PHOSPHORYLATION OCCURS CONCOMITANT WITH THE
CC      CATALYTIC STEPS OF SPLICING.
CC      -1- SIMILARITY: Belongs to the SF3B1 family.
CC      -1- SIMILARITY: Contains ? HEAT repeats.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      EMBL; AF054284; AAC97189.1; -

```

```

DR      EMBL; AF070540; AAC28633.1; -
DR      Genew; HGNC:10768; SF3B1.
DR      GK; Q75533; -
DR      MIM; 605590; -
DR      GO; GO:0005681; C:spliceosome complex; NAS.
DR      GO; GO:0008248; F:pre-mRNA splicing factor activity; NAS.
DR      GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; NAS.
DR      InterPro; IPR008938; ARM.
KW      Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;
KW      Phosphorylation; Repeat.
SQ      PHOSPHORYLATION; Repeat.
SQ      SEQUENCE 1304 AA; 145814 MW; 40DB21A6209165A7 CRC64;

Query Match      100.0%; Score 35; DB 1; Length 1304;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDTG 6
DB      269 GRGDTG 274

RESULT 3
S3B1_MOUSE
ID      S3B1_MOUSE      STANDARD;      PRT; 1304 AA.
AC      Q99NB9;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Splicing factor 3B subunit 1 (Spliceosome associated protein 155) (SAP
DE      155) (SF3B155) (Pre-mRNA splicing factor SF3B 155 kDa subunit).
GN      SF3B1 OR SAPI55.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RS      SEQUENCE FROM N.A.
RE      TISSUE=Ovary;
RA      MEDLINE=21189985; PubMed=11252167;
RA      Isono K., Abe K., Tomaru Y., Okazaki Y., Hayashizaki Y., Koseki H.;
RT      "Molecular cloning, genetic mapping, and expression of the mouse SF3b1
RT      (SAPI55) gene for the U2snRNP component of spliceosome.";
RL      Mamm. Genome 12:192-198 (2001).
CC      -1- FUNCTION: Subunit of the splicing factor SF3B required for 'A'
CC      complex assembly formed by the stable binding of U2 snRNP to the
CC      branchpoint sequence (bps) in pre-mRNA. Sequence independent
CC      binding of SF3A/SF3B complex upstream of the branch site is
CC      essential, it may anchor U2 snRNP to the pre-mRNA. May also be
CC      involved in the assembly of the 'E' complex. Belongs also to the
CC      minor U2-dependent spliceosome, which is involved in the splicing
CC      of rare class of nuclear pre-mRNA intron (By similarity).
CC      -1- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF
CC      FOUR SUBUNITS; SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,
CC      SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A
CC      12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS
CC      COMPLEX (U2 snRNP). SF3B1 INTERACTS DIRECTLY WITH THE SPLICING
CC      FACTOR U2AF (By similarity).
CC      -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC      -1- TISSUE SPECIFICITY: Ubiquitous.
CC      -1- PTM: PHOSPHORYLATED. PHOSPHORYLATION OCCURS CONCOMITANT WITH THE
CC      CATALYTIC STEPS OF SPLICING (By similarity).
CC      -1- SIMILARITY: Belongs to the SF3B1 family.
CC      -1- SIMILARITY: Contains ? HEAT repeats.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      EMBL; AB037890; BAB40140.1; -

```



DR MGD; XGI:1932339; Sflb1.  
 DR InterPro; I2R008938; ARM.  
 KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;  
 KW Phosphorylation; Repeat.  
 SQ SEQUENCE 1304 AA; 145816 MW; 12F051757D2A2DEE CRC64;

Query Match 100.0%; Score 35; DB 1; Length 1304;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6  
 DB 269 GRGDTTP 274

RESULT 4  
 ID S3B1 XENLA STANDARD; PRT; 1307 AA.  
 AC 057683;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Splicing factor 3B subunit 1 (Spliceosome associated protein 155) (SAP  
 155) (S23B155) (Pre-mRNA splicing factor SF3b 155 kDa subunit) 1146  
 DE kDa nuclear protein).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98099695; PubMed=9436997;  
 RA Schmidt-Zachmann M.S., Knecht S., Kraemer A.;  
 RT "Molecular characterization of a novel, widespread nuclear protein  
 that colocalizes with spliceosome components.";  
 RL Mol. Biol. Cell 9:143-150(1998).  
 CC -!- FUNCTION: Subunit of the splicing factor SF3B required for 'A'  
 CC complex assembly formed by the stable binding of U2 snRNP to the  
 CC branchpoint sequence (bps) in pre-mRNA. Sequence independent  
 CC binding of SF3A/SF3B complex upstream of the branch site is  
 CC essential, it may anchor U2 snRNP to the pre-mRNA. May also be  
 CC involved in the assembly of the '2' complex. Belongs also to the  
 CC minor U2-dependent spliceosome, which is involved in the splicing  
 CC of rare class of nuclear pre-mRNA intron (By similarity).  
 CC -!- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF  
 CC FOUR SUBUNITS: SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,  
 CC SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A  
 CC 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS  
 CC COMPLEX (U2 snRNP) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Belongs to the SF3B1 family.  
 CC -!- SIMILARITY: Contains ? HEAT repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y08997; CAA70201.1; -.  
 DR PIR; T30887; T30887.  
 DR InterPro; IPR008938; ARM.  
 KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;  
 KW Repeat.  
 SQ SEQUENCE 1307 AA; 146213 MW; 6B226D5F036C69B3 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 1307;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6  
 DB 272 GRGDTTP 277

RESULT 5  
 ID RL7A CHICK STANDARD; PRT; 265 AA.  
 AC F32429;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE 60S ribosomal protein L7a.  
 GN RL7A OR SURF-3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Liver;  
 RX MEDLINE=92335297; PubMed=1631131;  
 RA Colombo P., Yon J., Garson K., Fried M.;  
 RT "Conservation of the organization of five tightly clustered genes  
 RT over 600 million years of divergent evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6358-6362(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Liver;  
 RX MEDLINE=94100270; PubMed=8274530;  
 RA Maeda N., Kenmochi N., Tanaka T.;  
 RT "The complete nucleotide sequence of chicken ribosomal protein L7a  
 RT gene and the multiple factor binding sites in its 5'-flanking  
 RT region.";  
 RL Biochimie 75:785-790(1993).  
 CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X62640; CAA44506.1; -.  
 DR EMBL; D14522; BAA03395.1; -.  
 DR PIR; A46032; S18159.  
 DR InterPro; IPR004038; Ribosomal L7A.  
 DR InterPro; IPR004037; Ribosomal L7Ae.  
 DR Pfam; PF01248; Ribosomal L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
 KW Ribosomal protein.  
 FT INIT MET 0 0 BY SIMILARITY.  
 SQ SEQUENCE 265 AA; 29868 MW; 07E312C639BB5655 CRC64;

Query Match 91.4%; Score 32; DB 1; Length 265;  
 Best Local Similarity 83.3%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6  
 DB 123 GRGDTTP 128

RESULT 6  
 ID MB22 ARATH STANDARD; PRT; 458 AA.  
 AC O80950;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosinase binding protein-like At2g39310.  
GN AT2G39310 OR T16B24.5.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.-J., Rensing C.M., Koo H.L.,  
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,  
RA Tallon L.J., Gill J.E., Adams M.D., Carrara A.J., Creasy T.H.,  
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
RA Venter J.C.;  
RT \*Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana.\*  
RL Nature 402:761-768(1999).  
CC -1- SIMILARITY: Belongs to the jacalin lectin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AC004697; AAC28979.1;  
CC PIR; T02571; T02571.  
CC HGSP; P18674; L10T.  
CC InterPro; IPR001229; Jacalin\_lectin.  
CC Pfam; PF01419; Jacalin; 3.  
CC Lentin; Repeat; Multigene family.  
KW Lectin; Repeat; Multigene family.  
SQ SEQUENCE 458 AA; 50463 MW; E501A410563EAPF8 CRC64;  
  
Query Match 91.4%; Score 32; DB 1; Length 458;  
Best Local Similarity 83.3%; Pred. No. 19;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDTP 6  
DB 214 GKGDTP 219  
  
RESULT 7  
PAT2 CAEEL STANDARD; PRT; 1226 AA.  
ID PAT2 CAEEL STANDARD; PRT; 1226 AA.  
AC P34436;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha pat-2 precursor.  
GN PAT-2 OR F54F2.1.  
CS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peleciderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2.  
RX MEDLINE=94150718; PubMed=73061393;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dearn S., Du Z., Durbin R., Favell A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kerstew J., Kirsten J., Laissner N.,  
RA Latreille P., Lightwing J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wooldman P.;  
RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.\*  
RL Nature 368:32-38(1994).  
CC -1- FUNCTION: POSSIBLE ROLE IN CELL-CELL INTERACTIONS (BY SIMILARITY).  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA PAT-2  
CC ASSOCIATES WITH BETA PAT-3.  
CC  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L23645; AAK26134.1;  
CC PIR; S44824; S44824.  
CC HGSP; P06756; L1V2.  
CC WormPep; F54F2.1; CE00194.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC Pfam; PF01839; FG-GAP; 4.  
CC Pfam; PF03037; Integrin\_A; 1.  
CC PRINTS; PR01185; INTEGRIN.  
CC SMART; SM00191; Int alpha; 5.  
CC PROSITE; PS00242; INTEGRIN ALPHA; 1.  
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat.  
FT CHAIN 1 25  
FT SIGNAL 26 1226  
FT DOMAIN 26 1154  
FT TRANSMEM 1155 1177  
FT DOMAIN 1178 1226  
FT REPEAT 40 103  
FT REPEAT 120 172  
FT REPEAT 189 243  
FT REPEAT 244 297  
FT REPEAT 300 372  
FT REPEAT 373 433  
FT REPEAT 437 485  
FT CARBOHYD 108 108  
FT CARBOHYD 228 228  
FT CARBOHYD 290 290  
FT CARBOHYD 608 608  
FT CARBOHYD 679 679  
FT CARBOHYD 775 775  
FT CARBOHYD 819 819  
SQ SEQUENCE 1226 AA; 135939 MW; B9169AD75B88901D CRC64;  
  
Query Match 91.4%; Score 32; DB 1; Length 1226;  
Best Local Similarity 83.3%; Pred. No. 58;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GRGDTP 6  
DB 519 GKGDTP 524  
  
RESULT 8  
YIK3 YEAST STANDARD; PRT; 425 AA.  
ID YIK3 YEAST STANDARD; PRT; 425 AA.  
AC P40487.  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 48.3 kDa protein in MOB1-SGA1 intergenic region.

```

GN Y11103W.
OS Saccharomyces cerevisiae 'Baker's yeast'.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churche C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Iye G.,
RA Moulis S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
CC -!- SIMILARITY: TO YEAST DIPHTHERIA TOXIN RESISTANCE PROTEIN 2 (DPH2),
CC TO S.POMBE SPAC134.15C, SPBC3B8.05, TO C.ELEGANS C09G5.2, C14B1.5
CC AND M.JANKASCHII MJ0483.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z38.25; CAAB6277.1; -.
CC PIR; S48469; S48469.
CC GERMOnline; 139638; -.
CC SGD; S0001365; Y11103W.
CC InterPro; IPR002728; Diphthamide syn.
CC Pfam; PF01866; Diphthamide syn; 1.
CC ProDom; PD004399; Diphthamide syn; 1.
CC TIGSFAM; TIGR00322; diphth2_R; 1.
CC Hypothetical protein.
CC KW Hypothetical protein.
CC
CC QUERY MATCH 88.6%; Score 31; DB 1; Length 425;
CC Best Local Similarity 83.3%; Pred. No. 30;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 GRGDTP 6
CC DB 415 GRGETP 420
CC
CC RESULT 9
CC Y863_S3NY3
CC ID Y863_S3NY3 STANDARD; PRT; 463 AA.
CC AC P73754;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Hypothetical protein slr0863.
CC GN SLR0863.
CC OS Synechocystis sp. (strain PCC 6803).
CC OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
CC OX NCBI_TaxID=1148;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=97061201; PubMed=8905231;
CC RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
CC Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
CC Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
CC Yamada M., Yasuda M., Tabata S.;
CC "Sequence analysis of the genome of the unicellular cyanobacterium
CC Synechocystis sp. strain PCC6803. II. Sequence determination of the
CC entire genome and assignment of potential protein-coding regions.";
CC DNA Res. 3:109-136(1996).
CC -!- SIMILARITY: Belongs to the tldD/pmbA family.
CC

```

```

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90909; SAA17806.1; -.
CC PIR; S74845; S74845.
CC InterPro; IPR002510; Peptidase_U62.
CC Pfam; PF01523; PmbA tldD; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 463 AA; 50369 MW; 22017307C7C6A32D CRC64;
CC
CC QUERY MATCH 88.6%; Score 31; DB 1; Length 463;
CC Best Local Similarity 83.3%; Pred. No. 33;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 GRGDTP 6
CC DB 352 GRGETP 357
CC
CC RESULT 10
CC F1NC_CHICK
CC ID F1NC_CHICK STANDARD; PRT; 1256 AA.
CC AC P11722; Q90921;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Fibronectin (FN) (Fragments).
CC GN FNI.
CC OS Gallus gallus (Chicken).
CC OC Archoyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
CC OX NCBI_TaxID=9031;
CC RN [1]
CC RP SEQUENCE OF 1-50 FROM N.A.
CC RX MEDLINE=83117850; PubMed=6572007;
CC RA Hirano H., Yamada Y., Sullivan M., de Crombrughe B., Pastan I.,
CC Yamada K.M.;
CC "Isolation of genomic DNA clones spanning the entire fibronectin
CC gene.";
CC Proc. Natl. Acad. Sci. U.S.A. 80:46-50(1983).
CC RN [2]
CC RP SEQUENCE OF 51-1256 FROM N.A.
CC RC STRAIN=White leghorn;
CC RA Norton P.A.;
CC RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC RN [3]
CC RP SEQUENCE OF 227-415 FROM N.A.
CC RX MEDLINE=96183658; PubMed=8603103;
CC RA Gehris A.L., Brandli D.W., Lewis S.D., Bennett V.D.;
CC "The exon encoding the fibronectin type III-9 repeat is
CC constitutively included in the mRNA from chick limb mesenchyme and
CC cartilage.";
CC Biochim. Biophys. Acta 1311:5-12(1996).
CC RN [4]
CC RP SEQUENCE OF 327-599 FROM N.A.
CC RX MEDLINE=88050950; PubMed=2823899;
CC RA Kubomura S., Osara M., Karasaki Y., Taniguchi H., Gotoh S.,
CC Tsuda T., Higashi K., Ohsato K., Hiarno H.;
CC "Genetic analysis of the cell binding domain region of the chicken
CC fibronectin gene.";
CC Biochim. Biophys. Acta 910:171-181(1987).
CC RN [5]
CC RP SEQUENCE OF 413-1256 FROM N.A.
CC RX MEDLINE=88142820; PubMed=2830487;
CC RA Norton P.A., Hynes R.O.;
CC "Alternative splicing of chicken fibronectin in embryos and in normal

```

RT and transformed cells.":

RL Mol. Cell. Biol. 7:4297-4307(1987).

CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds

CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins

CC are involved in cell adhesion, cell motility, opsonization, wound

CC healing, and maintenance of cell shape.

CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED

CC VARIANTS. CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;

CC TO A LESSER EXTENT HOMODIMERS.

CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=1;

CC Comment=A number of isoforms are produced. Each of the "extra

CC domain" and the connecting strand 3 are present in some forms of

CC fibronectin and absent in others;

CC Name=1;

CC IsoId=PII1722-1; Sequence=displayed;

CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted

CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric

CC form), made by fibroblasts, epithelial and other cell types, is

CC deposited as fibrils in the extracellular matrix.

CC -!- PTM: Sulfated (By similarity).

CC -!- SIMILARITY: Contains at least 2 fibronectin type I domains.

CC -!- SIMILARITY: Contains at least 8 fibronectin type III domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; V00432; CAA23714.1; -

CC EMBL; U21327; AAA73566.1; -

CC EMBL; X06533; CAA29781.1; -

CC EMBL; M26186; AAA48772.1; ALT\_SEQ.

CC EMBL; U20386; AAB01062.1; -

CC PIR; A28512; A28512.

CC PIR; A29355; A29355.

CC PIR; S71465; S71465.

CC HSSP; P02751; 1FNH.

CC InterPro; IPR000083; Fibnctn1.

CC InterPro; IPR008957; FN III-like.

CC InterPro; IPR003961; FN-III.

CC InterPro; IPR000562; FN-Type II.

CC InterPro; IPR003962; FN-III\_subd.

CC PRINTS; PR00014; FNTYPEIII.

CC SMART; SM00058; FN1; 3.

CC SMART; SM00060; FN3; 9.

CC PROSITE; PS00023; FIBRONECTIN 2; PARTIAL.

CC PROSITE; PS01253; FIBRONECTIN 1; 2.

CC KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;

CC Repeat; Sulfation; Alternative splicing.

CC FT NON TER 1 1

CC FT 50 51

CC FT 236 509

CC FT 690 961

CC FT 1153 1226

CC FT 327 415

CC FT 416 509

CC FT 510 599

CC FT 600 689

CC FT 690 781

CC FT 782 871

CC FT 872 961

CC FT 1082 1082

CC FT 1083 1152

CC FT 1174 1218

CC FT 1219 >1256

CC FT 493 495

CC FT 1176 1205

CC FT 1203 1215

CC BY SIMILARITY.

FT DISULFID 1221 1248 BY SIMILARITY.

FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT CARBOHYD 1034 1034 O-LINKED (GALNAC. . .) (BY SIMILARITY).

FT CARBOHYD 1035 1035 O-LINKED (GALNAC. . .) (BY SIMILARITY).

FT CONFLICT 516 516 Q -> P (IN REF. 4).

FT CONFLICT 569 572 EGLQ -> QGLE (IN REF. 4).

FT NON TER 1256 1256

SQ SEQUENCE 1256 AA; 137435 MW; 345A4CA0E4D71D9B CRC64;

Query Match 88.6%; Score 31; DB 1; Length 1256;

Best Local Similarity 83.3%; Pred.No. 99;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 492 GRGDSF 497

RESULT 11

FINC\_PLEWA STANDARD; PRT; 1328 AA.

ID AC Q91283;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Fibronectin (FN) (Fragment).

OS Pleurodeles waltlii (Iberian ribbed newt).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;

OC Pleurodeles.

OX NCBI\_TaxID=8319;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94363379; PubMed=9081872;

RA Cavalier L., Riou J., Desmone D.N.;

RT "Amphibian Pleurodeles waltlii fibronectin: cDNA cloning and

RT developmental expression of spliced variants.";

RL Cell Adhes. Commun. 1:83-91(1993).

CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds

CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins

CC are involved in cell adhesion, cell motility, opsonization, wound

CC healing, and maintenance of cell shape (By similarity).

CC -!- SUBUNIT: DIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS,

CC CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS (BY

CC SIMILARITY).

CC -!- SIMILARITY: Contains at least 3 fibronectin type I domains.

CC -!- SIMILARITY: Contains at least 10 fibronectin type III domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X66813; CAA47292.1; -

CC HSSP; P02751; 1FNH.

CC InterPro; IPR006209; EGF-like.

CC InterPro; IPR000083; Fibnctn1.

CC InterPro; IPR008957; FN III-like.

CC InterPro; IPR003961; FN-III.

CC InterPro; IPR003962; FN-III\_subd.

CC Pfam; PF00039; fn1; 3.

CC Pfam; PF00041; fn3; 11.

CC PRINTS; PR00014; FNTYPEIII.

CC SMART; SM00058; FN1; 3.

CC SMART; SM00060; FN3; 9.

CC PROSITE; PS00022; EGF 1; 1.

CC PROSITE; PS01253; FIBRONECTIN 1; 1.

CC KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;

CC Repeat.

```

FT NON TER 1 1
FT DNA BIND <1 13
FT DOVAIN 203 477
FT DOVAIN 658 929
FT DOVAIN 1157 1288
FT DOVAIN <1 14
FT DOVAIN 15 110
FT DOVAIN 111 201
FT DOVAIN 202 292
FT DOVAIN 293 383
FT DOVAIN 384 477
FT DOVAIN 478 567
FT DOVAIN 568 657
FT DOVAIN 658 749
FT DOVAIN 750 838
FT DOVAIN 839 929
FT DOVAIN 930 1064
FT DOVAIN 1065 1143
FT DOVAIN 1155 1199
FT DOVAIN 1200 1243
FT DOVAIN 1245 1287
FT DISULFID 1157 1186
FT DISULFID 1184 1196
FT DISULFID 1202 1229
FT DISULFID 1227 1240
FT DISULFID 1247 1270
FT DISULFID 1268 1284
FT DISULFID 1320 1320
FT DISULFID 1324 1334
FT SITE 461 463
FT CARBOHYD 89 89
SQ SEQUENCE 1328 AA; 145037 MW; E31BF7965AID1E74 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 1328;
Best Local Similarity 83.3%; Pred. No. 1,1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRGDT 6
Db 460 GRGDSP 465

RESULT 12
FNC BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Fibronectin (FN).
GN FNI.
OS Bos taurus (Bovine).
OC Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX SEQUENCE.
RP MEDLINE=87054047; PubMed=3780752;
RA Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
RT "Complete primary structure of bovine plasma fibronectin.";
RL Eur. J. Biochem. 161:441-453 (1986).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=83117805; PubMed=6218503;
RA Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,
RA Sahl P., Sottrup-Jensen L., Magnusson S.;
RT "Partial primary structure of bovine plasma fibronectin: three types
RT of internal homology.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141 (1983).
RN [3]
RP SEQUENCE OF 2170-2265 FROM N.A.
RX MEDLINE=83221567; PubMed=6304699;

```

```

RA Kornblitt A.R., Vibe-Pedersen K., Baralle P.E.;
RT "Isolation and characterization of cDNA clones for human and bovine
RT fibronectins.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222 (1983).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins
CC are involved in cell adhesion, cell motility, opsonization, wound
CC healing, and maintenance of cell shape.
CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
CC TO A LESSER EXTENT HOMODIMERS.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced. Each of the "extra
CC domain" and the connecting strand 3 are present in some forms of
CC fibronectin and absent in others;
CC Name=1;
CC IsoId=P07589-1; Sequence-Displayed;
CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
CC forms), made by fibroblasts, epithelial and other cell types, is
CC deposited as fibrils in the extracellular matrix.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 15 fibronectin type III domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K00800; AAA30521.2; -.
CC PIR; A28452; FNBO.
CC DR HSSP; P02751; 2FN2.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR000083; Fibnctnl.
CC DR InterPro; IPR008957; FN III-like.
CC DR InterPro; IPR003961; FN_III.
CC DR InterPro; IPR000562; FN_Type_II.
CC DR InterPro; IPR003962; FNIII_subd.
CC Pfam; PF00039; fn1; 12.
CC Pfam; PF00040; fn2; 2.
CC Pfam; PF00041; fn3; 15.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00014; FNTYPEIII.
CC ProDom; PD000995; FN_Type_II; 2.
CC SMART; SM00058; FN1; 12.
CC SMART; SM00059; FN2; 2.
CC SMART; SM00060; FN3; 14.
CC PROSITE; PS00022; EGF 1; 2.
CC PROSITE; PS00023; FIBONECTIN 2; 2.
CC PROSITE; PS01253; FIBONECTIN 1; 12.
CC Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
CC Sulfation; Cell adhesion; Repeat; Alternative splicing;
CC Pyridolone carboxylic acid.
CC MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
CC FT DOVAIN 21 241 FIBRI- AND HEPARIN-BINDING 1.
CC FT DOVAIN 277 577 COLLAGEN-BINDING.
CC FT DOVAIN 876 1141 CELL-ATTACHMENT.
CC FT DOVAIN 1236 1509 HEPARIN-BINDING 2.
CC FT DOVAIN 1600 1870 FIBRI- AND HEPARIN-BINDING 2.
CC FT DOVAIN 1991 2216 FIBONECTIN TYPE-I 1.
CC FT DOVAIN 19 59 FIBONECTIN TYPE-I 2.
CC FT DOVAIN 64 107 FIBONECTIN TYPE-I 3.
CC FT DOVAIN 108 151 FIBONECTIN TYPE-I 4.
CC FT DOVAIN 153 197 FIBONECTIN TYPE-I 5.
CC FT DOVAIN 198 242 FIBONECTIN TYPE-I 6.
CC FT DOVAIN 275 314

```

FT DOMAIN 314 373 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 374 438 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 437 480 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 485 527 FIBRONECTIN TYPE-I 8.  
 FT DOMAIN 528 571 FIBRONECTIN TYPE-I 9.  
 FT DOMAIN 578 669 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 688 778 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 779 874 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 875 964 FIBRONECTIN TYPE-III 4.  
 FT DOMAIN 965 1054 FIBRONECTIN TYPE-III 5.  
 FT DOMAIN 1055 1141 FIBRONECTIN TYPE-III 6.  
 FT DOMAIN 1142 1234 FIBRONECTIN TYPE-III 7.  
 FT DOMAIN 1235 1325 FIBRONECTIN TYPE-III 8.  
 FT DOMAIN 1326 1415 FIBRONECTIN TYPE-III 9.  
 FT DOMAIN 1416 1509 FIBRONECTIN TYPE-III 10.  
 FT DOMAIN 1510 1599 FIBRONECTIN TYPE-III 11. (EXTRA DOMAIN).  
 FT DOMAIN 1600 1691 FIBRONECTIN TYPE-III 12.  
 FT DOMAIN 1692 1780 FIBRONECTIN TYPE-III 13.  
 FT DOMAIN 1781 1870 FIBRONECTIN TYPE-III 14.  
 FT DOMAIN 1871 1990 CONNECTING STRAND 3 (CS-3) (V REGION).  
 FT DOMAIN 1982 2061 FIBRONECTIN TYPE-III 15.  
 FT DOMAIN 2083 2127 FIBRONECTIN TYPE-I 10.  
 FT DOMAIN 2128 2170 FIBRONECTIN TYPE-I 11.  
 FT DOMAIN 2172 2215 FIBRONECTIN TYPE-I 12.  
 FT SITE 1493 1495 CELL ATTACHMENT SITE.  
 FT DISULFID 21 47  
 FT DISULFID 45 56  
 FT DISULFID 66 94  
 FT DISULFID 92 104  
 FT DISULFID 110 138  
 FT DISULFID 136 148  
 FT DISULFID 155 184  
 FT DISULFID 182 194  
 FT DISULFID 200 229  
 FT DISULFID 227 239  
 FT DISULFID 277 304  
 FT DISULFID 302 311  
 FT DISULFID 329 355  
 FT DISULFID 343 370  
 FT DISULFID 389 415  
 FT DISULFID 403 430  
 FT DISULFID 439 467  
 FT DISULFID 465 477  
 FT DISULFID 512 524  
 FT DISULFID 530 558  
 FT DISULFID 556 568  
 FT DISULFID 2085 2114  
 FT DISULFID 2112 2124  
 FT DISULFID 2130 2157  
 FT DISULFID 2155 2167  
 FT DISULFID 2174 2200  
 FT DISULFID 2198 2209  
 FT DISULFID 2246 2246  
 FT DISULFID 2250 2250  
 FT MOD RES 845 845  
 FT MOD RES 850 850  
 FT CARBOHYD 399 399  
 FT CARBOHYD 497 497  
 FT CARBOHYD 511 511  
 FT CARBOHYD 846 846  
 FT CARBOHYD 976 976  
 FT CARBOHYD 1213 1213  
 FT CARBOHYD 1987 1987  
 FT CARBOHYD 1943 1943  
 FT CARBOHYD 1944 1944  
 FT MOD RES 2263 2263  
 SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;  
 Query Match 88.68; Score 31; DB 1; Length 2265;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6  
 Db 1492 GRGDSF 1497  
 RESULT 13  
 ID FINE HUMAN STANDARD; PRT; 2386 AA.  
 AC P02751; Q95609; Q14312; Q14325; Q14326; Q86T27; Q81V18;  
 AC Q96KP7; Q96KP8; Q96KP9; Q9H1B8; Q9HAP3; Q9UMK2;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Fibrinectin precursor (FN) (Cold-insoluble globulin) (CIG).  
 GN FN1 OR FN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=21600194; PubMed=11737888;  
 RA Schor S.L., Schor A.M.;  
 RT "Phenotypic and genetic alterations in mammary stroma: implications  
 RT for tumour progression.";  
 RL Breast Cancer Res. 3:373-379(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 3; 7 AND 10).  
 RC TISSUE=Cervix;  
 RA Ansgorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
 RA Mewes H.-W., Weil B., Amid C., Osanger A., Fobo G., Han M.,  
 RA Wiemann S.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-38 FROM N.A.  
 RX MEDLINE=87030890; PubMed=3770189;  
 RA Gutman A., Yamada K.M., Kornblitt A.R.;  
 RT "Human fibronectin is synthesized as a pre-propolyptide.";  
 RL FEBS Lett. 207:145-148(1986).  
 RN [4]  
 RP SEQUENCE OF 1-43 FROM N.A.  
 RX MEDLINE=87175578; PubMed=3031656;  
 RA Dean D.C., Bowlus C.L., Bourgeois S.;  
 RT "Cloning and analysis of the promoter region of the human fibronectin  
 RT gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:1876-1880(1987).  
 RN [5]  
 RP SEQUENCE OF 28-2386 FROM N.A. (ISOFORM 3).  
 RX MEDLINE=85284965; PubMed=2992939;  
 RA Kornblitt A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E.;  
 RT "Primary structure of human fibronectin: differential splicing may  
 RT generate at least 10 polypeptides from a single gene.";  
 RL EMBO J. 4:1755-1759(1985).  
 RN [6]  
 RP SEQUENCE OF 103-481 AND 2228-2386 FROM N.A. (ISOFORMS 1; 3; 8 AND 9).  
 RC TISSUE=Peripheral blood T-cell, and Umbilical vein endothelial cells;  
 RA Godfrey H.P., Ebrahim A.A.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 973-2386 FROM N.A. (ISOFORM 3).  
 RX MEDLINE=84272258; PubMed=6462919;  
 RA Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;  
 RT "Human fibronectin: cell specific alternative mRNA splicing generates  
 RT polypeptide chains differing in the number of internal repeats.";  
 RL Nucleic Acids Res. 12:5853-5868(1984).  
 RN [8]  
 RP SEQUENCE OF 1232-1782 FROM N.A. (ISOFORM 7).  
 RX MEDLINE=88233940; PubMed=3375063;  
 RA Paoletti G., Henschliffe C., Sebastiao G., Baralle F.E.;  
 RT "Sequence analysis and in vivo expression show that alternative  
 RT splicing of ED-B and ED-A regions of the human fibronectin gene are  
 RT independent events.";  
 RL Nucleic Acids Res. 16:3545-3557(1988).

RN [9] SEQUENCE OF 1257-1365 FROM N.A. (ISOFORM 11).  
 RP MEDLINE=88041070; PubMed=3478690;  
 RX Gutman A., Kornblitt A.R.;  
 RA "Identification of a third region of cell-specific alternative  
 RT splicing in human fibronectin mRNA."; 7182(1987).  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7179-7182(1987).  
 RN [10]  
 RN SEQUENCE OF 1441-1548.  
 RP MEDLINE=82265604; PubMed=7050098;  
 RX Piersbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;  
 RA "The cell attachment domain of fibronectin. Determination of the  
 RT primary structure."; 257(1982).  
 RL J. Biol. Chem. 257:9593-9597(1982).  
 RN [11]  
 RN SEQUENCE OF 1448-1540 FROM N.A.  
 RP MEDLINE=83290929; PubMed=6688418;  
 RX Oldberg A., Linney E., Ruoslahti E.;  
 RA "Molecular cloning and nucleotide sequence of a cDNA clone coding for  
 RT the cell attachment domain in human fibronectin."; 258(1983).  
 RL J. Biol. Chem. 258:10193-10196(1983).  
 RN [12]  
 RN SEQUENCE OF 1448-1540 FROM N.A.  
 RP MEDLINE=86111901; PubMed=3003095;  
 RX Oldberg A., Ruoslahti E.;  
 RA "Evolution of the fibronectin gene. Exon structure of cell attachment  
 RT domain."; 261(1986).  
 RL J. Biol. Chem. 261:2113-2116(1986).  
 RN [13]  
 RN SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).  
 RP MEDLINE=85280409; PubMed=2992573;  
 RX Bernard M.P., Kolbe M., Weil D., Chu M.-L.;  
 RA "Human cellular fibronectin: comparison of the carboxyl-terminal  
 RT portion with rat identifies primary structural domains separated by  
 RT hypervariable regions."; 24(1985).  
 RL Biochemistry 24:2698-2704(1985).  
 RN [14]  
 RN SEQUENCE OF 1712-1739 FROM N.A.  
 RP MEDLINE=87026578; PubMed=3021206;  
 RX Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;  
 RA "Human liver fibronectin complementary DNAs: identification of two  
 RT different messenger RNAs possibly encoding the alpha and beta  
 RT subunits of plasma fibronectin."; 25:4936-4941(1986).  
 RL Biochemistry 25:4936-4941(1986).  
 RN [15]  
 RN SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).  
 RP TISSUE=Cartilage;  
 RX MEDLINE=22126816; PubMed=12127832;  
 RA Parker A.E., Boutell J., Carr A., Maciewicz R.A.;  
 RT "Novel cartilage-specific splice variants of fibronectin."; 10:528-534(2002).  
 RL Osteoarthritis Cartilage 10:528-534(2002).  
 RN [16]  
 RN SEQUENCE OF 32-290.  
 RP MEDLINE=84032463; PubMed=6630202;  
 RX Garcia-Pardo A., Pearlstein E., Frangione B.;  
 RA "Primary structure of human plasma fibronectin. The 28,000-dalton  
 RT NH2-terminal domain."; 258:12670-12674(1983).  
 RL J. Biol. Chem. 258:12670-12674(1983).  
 RN [17]  
 RN SEQUENCE OF 309-608. AND COLLAGEN-BINDING.  
 RP MEDLINE=87080265; PubMed=3024962;  
 RX Owens R.J., Baralle F.E.;  
 RA "Mapping the collagen-binding site of human fibronectin by expression  
 RT in Escherichia coli."; 5:2825-2830(1986).  
 RL EMBO J. 5:2825-2830(1986).  
 RN [18]  
 RN SULFATION.  
 RP MEDLINE=86042625; PubMed=2414772;  
 RX Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
 RA "Tyrosine sulfation of proteins from the human hepatoma cell line  
 RT HepG2."; 92:7160-7164(1985).  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7160-7164(1985).  
 RN [19]

RP O-GLYCOSYLATION OF THR-2064.  
 RX MEDLINE=91190085; PubMed=2012601;  
 RA Tresselt T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,  
 RA Shively J.E., Pande H.;  
 RT "Human plasma fibronectin. Demonstration of structural differences  
 RT between the A- and B-chains in the III CS region."; 274:731-738(1991).  
 RL Biochem. J. 274:731-738(1991).  
 RN [20]  
 RN FBLNI-BINDING SITE.  
 RP MEDLINE=93015879; PubMed=1400330;  
 RX Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,  
 RA Argaves W.S.;  
 RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding  
 RT region of fibronectin."; 267:20120-20125(1992).  
 RL J. Biol. Chem. 267:20120-20125(1992).  
 RN [21]  
 RN CHARACTERIZATION OF FIBRIN-BINDING SITE 1.  
 RP MEDLINE=95081151; PubMed=7989369;  
 RX Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;  
 RA "Further characterization of the NH2-terminal fibrin-binding site on  
 RT fibronectin."; 269:31938-31945(1994).  
 RL J. Biol. Chem. 269:31938-31945(1994).  
 RN [22]  
 RN STRUCTURE BY NMR OF 1447-1540.  
 RP MEDLINE=92162710; PubMed=1311202;  
 RX Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,  
 RA Campbell I.D.;  
 RT "1H NMR assignment and secondary structure of the cell adhesion type  
 RT III module of fibronectin."; 31:2068-2073(1992).  
 RL Biochemistry 31:2068-2073(1992).  
 RN [23]  
 RN STRUCTURE BY NMR OF 1447-1540.  
 RP MEDLINE=93046665; PubMed=1423622;  
 RX Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;  
 RA "The three-dimensional structure of the tenth type III module of  
 RT fibronectin: an insight into RGD-mediated interactions."; 71:671-678(1992).  
 RL Cell 71:671-678(1992).  
 RN [24]  
 RN STRUCTURE BY NMR OF 182-275.  
 RP MEDLINE=94141923; PubMed=8308892;  
 RX Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,  
 RA Campbell I.D.;  
 RT "Solution structure of a pair of fibronectin type 1 modules with  
 RT fibrin binding activity."; 235:1302-1311(1994).  
 RL J. Mol. Biol. 235:1302-1311(1994).  
 RN [25]  
 RN STRUCTURE BY NMR OF 32-92.  
 RP MEDLINE=96069779; PubMed=7583666;  
 RX Potts J.R., Phan I., Williams M.J., Campbell I.D.;  
 RA "High-resolution structural studies of the factor XIIIa crosslinking  
 RT site and the first type 1 module of fibronectin."; 2:946-950(1995).  
 RL Nat. Struct. Biol. 2:946-950(1995).  
 RN [26]  
 RN STRUCTURE BY NMR OF 406-464.  
 RP MEDLINE=98179558; PubMed=9514732;  
 RX Sticht H., Pickford A.R., Potts J.R., Campbell I.D.;  
 RA "Solution structure of the glycosylated second type 2 module of  
 RT fibronectin."; 276:177-187(1998).  
 RL J. Mol. Biol. 276:177-187(1998).  
 RN [27]  
 RN STRUCTURE BY NMR OF EXTRA ED-B DOMAIN FROM ISOFORM 7.

Query Match 88.6%; Score 31; DB 1; Length 2386;  
 Best Local Similarity 83.3%; Pred. No. 2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6

Db 1523 GRGDSP 1528

RESULT 14  
 FINE\_MOUSE

Copie V., Tomita Y., Akiyama S.K, Acta S., Yamada K.M., Venable R.M.,  
Pastor R.W., Krueger S., Torchia D.A.;  
"Solution structure and dynamics of linked cell attachment modules of  
mouse fibronectin containing the RGD and synergy regions: comparison  
with the human fibronectin crystal structure.";   
J. Mol. Biol. 277:663-682(1998).  
[8]  
DOWN-REGULATION BY GLUCOCORTICOIDS.  
MEDLINE=21609963; PubMed=11737251;  
Gu Y.-C., Tals J.F., Gulberg D., Timpl R., Ekblom M.;  
"Glucocorticoids down-regulate the extracellular matrix proteins  
fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma";  
Eur. J. Haematol. 67:176-184(2001).  
-!- FUNCTION: Fibronectins bind cell surfaces and various compounds  
including collagen, fibrin, heparin, DNA, and actin. Fibronectins  
are involved in cell adhesion, cell motility, opsonization, wound  
healing, and maintenance of cell shape.  
-!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced  
variants, connected by 2 disulfide bonds near the carboxyl end;  
to a lesser extent homodimers. Interacts with FBLN1 (By  
similarity).  
-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
-!- ALTERNATIVE PRODUCTS:  
Comment=-A number of isoforms are produced. Each of the "extra  
domain" and the connecting strand 3 are present in some forms of  
fibronectin and absent in others;  
Name=1;  
IsoId=PL11276-1; Sequence=Displayed;  
-!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted  
by hepatocytes. Cellular FN (dimeric or cross-linked multimeric  
forms), made by fibroblasts, epithelial and other cell types, is  
deposited as fibrils in the extracellular matrix.  
-!- INDUCTION: Glucocorticoids suppressed mRNA expression and protein  
synthesis.  
-!- PTM: Sulfated (By similarity).  
-!- SIMILARITY: Contains 12 fibronectin type I domains.  
-!- SIMILARITY: Contains 2 fibronectin type II domains.  
-!- SIMILARITY: Contains 17 fibronectin type III domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; BC051082; AAH51082.1; -  
EMBL; Z22729; CAA80422.1; -  
EMBL; X82402; CAA577956.1; -  
EMBL; X93167; CAA63654.1; -  
EMBL; M18194; AAA37636.1; -  
EMBL; S45680; AAB23491.1; -  
PIR; A49173; A49173.  
PIR; I48349; I48349.  
PDB; IMFN; 29-APR-98.  
PDB; 2MFN; 29-APR-98.  
PDB; MGI; 95566; Fnl.  
GO; GO:007155; P:cell adhesion; IDA.  
InterPro; IPR006209; EGF like.  
InterPro; IPR000083; Fibrinctnl.  
InterPro; IPR008957; FN III-like.  
InterPro; IPR003961; FN III.  
InterPro; IPR000562; FN type II.  
InterPro; IPR003962; FNIII\_subd.  
InterPro; IPR008924; WCR alpha\_beta\_C.  
PRINTS; PR00014; FNTYPEIII.  
SMART; SMC0058; FN1; 4.  
SMART; SMC0060; FN3; 12.  
PROSITE; PS00022; EGF\_1; 2.  
PROSITE; PS00023; FIBRONECTIN\_2; 2.  
PROSITE; PS01253; FIBRONECTIN\_1; 12.  
PROSITE; PS01253; FIBRONECTIN\_1; 12.



KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;  
 KW Sulfation; Cell adhesion; Repeat; Alternative splicing; Signal;  
 KW 3D-structure.

FT SIGNAL	1	32	BY SIMILARITY.
FT CHAIN	33	2477	FIBRONECTIN.
FT DOMAIN	53	273	FIBRIN- AND HEPARIN-BINDING 1.
FT DOMAIN	308	658	COLLAGEN-BINDING.
FT DNA_BIND	906	1171	CELL-ATTACHMENT.
FT DOMAIN	1357	1630	HEPARIN-BINDING 2.
FT DOMAIN	1811	2081	FIBRIN-BINDING 2.
FT DOMAIN	2236	2477	FIBRONECTIN TYPE-I 1.
FT DOMAIN	51	96	FIBRONECTIN TYPE-I 2.
FT DOMAIN	96	140	FIBRONECTIN TYPE-I 3.
FT DOMAIN	140	185	FIBRONECTIN TYPE-I 4.
FT DOMAIN	185	230	FIBRONECTIN TYPE-I 5.
FT DOMAIN	230	272	FIBRONECTIN TYPE-I 6.
FT DOMAIN	306	343	FIBRONECTIN TYPE-II 1.
FT DOMAIN	345	404	FIBRONECTIN TYPE-II 2.
FT DOMAIN	405	469	FIBRONECTIN TYPE-I 7.
FT DOMAIN	468	516	FIBRONECTIN TYPE-I 8.
FT DOMAIN	516	559	FIBRONECTIN TYPE-I 9.
FT DOMAIN	559	602	FIBRONECTIN TYPE-III 1.
FT DOMAIN	609	706	FIBRONECTIN TYPE-III 2.
FT DOMAIN	707	808	FIBRONECTIN TYPE-III 3.
FT DOMAIN	809	903	FIBRONECTIN TYPE-III 4.
FT DOMAIN	904	994	FIBRONECTIN TYPE-III 5.
FT DOMAIN	995	1084	FIBRONECTIN TYPE-III 6.
FT DOMAIN	1085	1172	FIBRONECTIN TYPE-III 7.
FT DOMAIN	1173	1264	FIBRONECTIN TYPE-III 8. (EXTRA DOMAIN 1).
FT DOMAIN	1265	1355	FIBRONECTIN TYPE-III 9.
FT DOMAIN	1356	1446	FIBRONECTIN TYPE-III 10.
FT DOMAIN	1447	1536	FIBRONECTIN TYPE-III 11.
FT DOMAIN	1537	1630	FIBRONECTIN TYPE-III 12.
FT DOMAIN	1631	1720	FIBRONECTIN TYPE-III 13. (EXTRA DOMAIN 2).
FT DOMAIN	1721	1810	FIBRONECTIN TYPE-III 14.
FT DOMAIN	1811	1902	FIBRONECTIN TYPE-III 15.
FT DOMAIN	1903	1991	FIBRONECTIN TYPE-III 16.
FT DOMAIN	1992	2081	CONNECTING STRAND 3 (CS-3) (V REGION).
FT DOMAIN	2082	2201	FIBRONECTIN TYPE-I 10.
FT DOMAIN	2202	2283	FIBRONECTIN TYPE-I 11.
FT DOMAIN	2284	2338	FIBRONECTIN TYPE-I 12.
FT DOMAIN	2339	2426	CELL ATTACHMENT SITE.
FT SITE	1614	1616	BY SIMILARITY.
FT SITE	2181	2183	BY SIMILARITY.
FT DISULFID	53	79	BY SIMILARITY.
FT DISULFID	77	88	BY SIMILARITY.
FT DISULFID	98	126	BY SIMILARITY.
FT DISULFID	124	136	BY SIMILARITY.
FT DISULFID	142	170	BY SIMILARITY.
FT DISULFID	168	180	BY SIMILARITY.
FT DISULFID	187	216	BY SIMILARITY.
FT DISULFID	214	226	BY SIMILARITY.
FT DISULFID	232	261	BY SIMILARITY.

Query Match 88.6%; Score 31; DB 1; Length 2477;  
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6  
 DB 1613 GRGDSP 1618

RESULT 15  
 FVNC RAT  
 IE FVNC RAT  
 AC P04937; STANDARD; PRT; 2477 AA.  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Fibrinectin precursor (FN).  
 GN FNL.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fischer; TISSUE=Liver;  
 RX MEDLINE=88054951; PubMed=2445560;  
 RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;  
 RT "Multiple sites of alternative splicing of the rat fibronectin gene transcript.";  
 RL EMBL J. 6:2573-2580(1987).  
 RN [2]  
 RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.  
 RC STRAIN=Fischer; TISSUE=Liver;  
 RX MEDLINE=88054950; PubMed=3119323;  
 RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;  
 RT "Organization of the fibronectin gene provides evidence for exon shuffling during evolution.";  
 RL EMBL J. 6:2565-2572(1987).  
 RN [3]  
 RP SEQUENCE OF 1586-2477 FROM N.A.  
 RX MEDLINE=84082067; PubMed=6317187;  
 RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;  
 RT "Three different fibronectin mRNAs arise by alternative splicing within the coding region.";  
 RL Cell 35:421-431(1983).  
 CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds including collagen, fibrin, heparin, DNA, and actin. Fibronectins are involved in cell adhesion, cell motility, opsonization, wound healing, and maintenance of cell shape.  
 CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS; TO A LESSER EXTENT HOMODIMERS.  
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=4;  
 CC Comment=Each of the "extra domain" and the connecting strand 3 are present in some forms of fibronectin and absent in others;  
 CC Name=1;  
 CC IsoId=P04937-1; Sequence=Displayed;  
 CC Name=2; Synonyms=FNII-13-less;  
 CC IsoId=P04937-2; Sequence=VSP\_003258;  
 CC Name=3; Synonyms=Lambda-RLF4-5;  
 CC IsoId=P04937-3; Sequence=VSP\_003259;  
 CC Name=4; Synonyms=Lambda-RLF6;  
 CC IsoId=P04937-4; Sequence=VSP\_003260;  
 CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted by hepatocytes. Cellular FN (dimeric or cross-linked multimeric forms), made by fibroblasts, epithelial and other cell types, is deposited as fibrils in the extracellular matrix.  
 CC -!- PTM: Sulfated (By similarity).  
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.  
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.  
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X15906; CAA34020.1; -  
 DR EMBL; L29191; AAA41166.1; -  
 DR EMBL; L00191; AAA41166.1; JOINED.  
 DR EMBL; L29191; AAA41167.1; -  
 DR EMBL; L00191; AAA41167.1; JOINED.  
 DR EMBL; L29191; AAA41168.1; -  
 DR EMBL; L00191; AAA41168.1; JOINED.  
 DR EMBL; X05831; CAA29278.1; -  
 DR EMBL; X05832; CAA29279.1; -

DR	EMBL; X05833; CAA29280.1; -.
DR	EMBL; X05834; CAA29281.1; -.
DR	PIR; S14428; S14428.
DR	HSSP; P02751; 1FBR.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; FibrinctnI.
DR	InterPro; IPR008957; FN III-like.
DR	InterPro; IPR003961; FN III.
DR	InterPro; PR000562; FN Type II.
DR	InterPro; IPR003962; FNII_subd.
DR	Pfam; PF00039; fn; 12.
DR	Pfam; PF00040; fn2; 2.
DR	Pfam; PF00041; fn3; 17.
DR	PRINTS; PR00013; FNTYPEII.
DR	PRINTS; PR00014; FNTYPEIII.
DR	ProDom; PD000955; FN_Type_II; 2.
DR	SMART; SM00058; FN1; 12.
DR	SMART; SM00059; FN2; 2.
DR	SMART; SM00060; FN3; 13.
DR	PROSITE; PS00022; EGF 1; 2.
DR	PROSITE; PS00023; FIBRONECTIN_2; 2.
DR	PROSITE; PS01253; FIBRONECTIN_1; 12.
KW	Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
KW	Sulfation; Cell adhesion; Repeat; Alternative splicing; Signal.
FT SIGNAL	1 32
FT CHAIN	33 2477 FIBRONECTIN.
FT DOMAIN	53 273 FIBRI- AND HEPARIN-BINDING 1.
FT DOMAIN	308 608 COLLAGEN-BINDING.
FT DNA_BIND	906 1171 CELL-ATTACHMENT.
FT DOMAIN	1357 163C HEPARIN-BINDING 2.
FT DOMAIN	1811 2081 FIBRI- BINDING 2.
FT DOMAIN	2296 2427 FIBRONECTIN TYPE-I 1.
FT DOMAIN	51 91 FIBRONECTIN TYPE-I 2.
FT DOMAIN	96 139 FIBRONECTIN TYPE-I 3.
FT DOMAIN	140 183 FIBRONECTIN TYPE-I 4.
FT DOMAIN	185 229 FIBRONECTIN TYPE-I 5.
FT DOMAIN	230 274 FIBRONECTIN TYPE-I 6.
FT DOMAIN	306 345 FIBRONECTIN TYPE-II 1.
FT DOMAIN	345 404 FIBRONECTIN TYPE-II 2.
FT DOMAIN	405 469 FIBRONECTIN TYPE-I 7.
FT DOMAIN	468 511 FIBRONECTIN TYPE-I 8.
FT DOMAIN	516 558 FIBRONECTIN TYPE-I 9.
FT DOMAIN	559 602 FIBRONECTIN TYPE-III 1.
FT DOMAIN	609 706 FIBRONECTIN TYPE-III 2.
FT DOMAIN	707 808 FIBRONECTIN TYPE-III 3.
FT DOMAIN	809 903 FIBRONECTIN TYPE-III 4.
FT DOMAIN	904 994 FIBRONECTIN TYPE-III 5.
FT DOMAIN	995 1084 FIBRONECTIN TYPE-III 6.
FT DOMAIN	1085 1172 FIBRONECTIN TYPE-III 7.
FT DOMAIN	1173 1264 FIBRONECTIN TYPE-III 8 {EXTRA DOMAIN 1}.
FT DOMAIN	1265 1355 FIBRONECTIN TYPE-III 9.
FT DOMAIN	1356 1446 FIBRONECTIN TYPE-III 10.
FT DOMAIN	1447 1536 FIBRONECTIN TYPE-III 11.
FT DOMAIN	1537 1630 FIBRONECTIN TYPE-III 12.
FT DOMAIN	1631 1720 FIBRONECTIN TYPE-III 13 {EXTRA DOMAIN 2}
FT DOMAIN	1721 1810 FIBRONECTIN TYPE-III 14.
FT DOMAIN	1811 1902 FIBRONECTIN TYPE-III 15.
FT DOMAIN	1903 1991 FIBRONECTIN TYPE-III 16.
FT DOMAIN	1992 2081 CONNECTING STRAND 3 {CS-3} {V REGION}.
FT DOMAIN	2082 2201 FIBRONECTIN TYPE-III 17.
FT DOMAIN	2202 2283 FIBRONECTIN TYPE-I 10.
FT DOMAIN	2294 2338 FIBRONECTIN TYPE-I 11.
FT DOMAIN	2339 2381 FIBRONECTIN TYPE-I 12.
FT DOMAIN	2383 2426 CELL ATTACHMENT SITE.
FT SITE	1614 1616 CELL ATTACHMENT SITE.
FT SITE	2181 2183 BY SIMILARITY.
FT DISULFD	53 79 BY SIMILARITY.
FT DISULFD	77 88 BY SIMILARITY.
FT DISULFD	98 126 BY SIMILARITY.
FT DISULFD	124 136 BY SIMILARITY.
FT DISULFD	142 170 BY SIMILARITY.
FT DISULFD	168 180 BY SIMILARITY.
FT DISULFD	-87 216 BY SIMILARITY.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 47.4 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-8  
Perfect score: 35  
Sequence: 1 GRGDTP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	35	100.0	354	5 Q9BL10	Q9BL10 bombyx mori
2	35	100.0	450	16 Q9AA55	Q9AA55 caulobacter
3	35	100.0	430	11 Q9CSK5	Q9CSK5 mus musculus
4	35	100.0	436	11 Q9ET34	Q9ET34 rattus norv
5	35	100.0	500	13 Q7ZTA4	Q7ZTA4 xenopus lae
6	35	100.0	575	16 Q89P25	Q89P25 brachyrihob
7	35	100.0	789	4 Q7Z497	Q7Z497 homo sapien
8	35	100.0	1209	5 Q8T8C0	Q8T8C0 bombyx mori
9	32	91.4	352	10 Q7XJ65	Q7XJ65 cucumis sat
10	32	91.4	458	10 Q8LBD2	Q8LBD2 arabidopsis
11	32	91.4	494	16 Q89LT8	Q89LT8 brachyrihob
12	32	91.4	606	16 Q8YKQ0	Q8YKQ0 anabaena sp
13	32	91.4	749	16 Q7VX39	Q7VX39 bordetella
14	32	91.4	1514	16 Q82BH0	Q82BH0 streptomyce
15	32	91.4	4246	5 Q7YUR0	Q7YUR0 trypanosoma
16	31	88.6	23	4 Q9UC00	Q9UC00 homo sapien

17	31	88.6	83	16 Q7WGA3	Q7WGA3 bordetella
18	31	88.6	83	16 Q7W4S7	Q7W4S7 bordetella
19	31	88.6	83	16 Q7WU91	Q7WU91 bordetella
20	31	88.6	86	2 Q9EZD2	Q9EZD2 rhizobium m
21	31	88.6	185	8 Q8HUM8	Q8HUM8 woodwardia
22	31	88.6	185	8 Q8HUM7	Q8HUM7 woodwardia
23	31	88.6	185	8 Q8HUM2	Q8HUM2 woodwardia
24	31	88.6	208	16 Q92K28	Q92K28 rhizobium m
25	31	88.6	217	16 Q7W3Z4	Q7W3Z4 bordetella
26	31	88.6	218	4 Q8N290	Q8N290 homo sapien
27	31	88.6	248	16 Q7WUW1	Q7WUW1 bordetella
28	31	88.6	255	16 Q7WFC6	Q7WFC6 bordetella
29	31	88.6	275	16 Q92PA3	Q92PA3 rhizobium m
30	31	88.6	293	6 Q9XSG0	Q9XSG0 oryctolagus
31	31	88.6	301	11 Q8CGC1	Q8CGC1 mus musculus
32	31	88.6	363	11 Q8CCB5	Q8CCB5 mus musculus
33	31	88.6	399	17 Q9HQ14	Q9HQ14 halobacteri
34	31	88.6	407	8 Q8SGB0	Q8SGB0 micropterus
35	31	88.6	438	2 Q9X2S7	Q9X2S7 vibrio chol
36	31	88.6	438	16 Q9XNL9	Q9XNL9 vibrio chol
37	31	88.6	441	17 Q9HQN2	Q9HQN2 halobacteri
38	31	88.6	515	16 Q8X6B0	Q8X6B0 escherichia
39	31	88.6	515	16 Q8FKH9	Q8FKH9 escherichia
40	31	88.6	562	4 Q8NC49	Q8NC49 homo sapien
41	31	88.6	589	13 Q9OYL3	Q9OYL3 brachydanio
42	31	88.6	589	13 Q7ZV14	Q7ZV14 brachydanio
43	31	88.6	606	5 Q18869	Q18869 caenorhabdi
44	31	88.6	618	17 Q8TYQ5	Q8TYQ5 methanopyru
45	31	88.6	659	10 Q8GWX7	Q8GWX7 arabidopsis

# ALIGNMENTS

RESULT 1  
Q9BL10 PRELIMINARY; PRT; 354 AA.  
AC Q9BL10; 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Inducible nitric oxide synthase-like protein (iNOS-LP)  
DE {Fragment}.  
GN iNOS-LP.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yang J., Tariat K., Yamakawa M.;  
RT "Inducible nitric oxide synthase-like protein (iNOS-LP) cDNA.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB017521; BAB33296.1; -.  
DR HSP; P29477; INOC  
DR GO; GO:0004517; F: nitric-oxide synthase activity; IEA.  
DR GO; GO:0006809; P: nitric oxide biosynthesis; IEA.  
DR InterPro; IPR004030; NO synthase.  
DR Pfam; PF02898; NO synthase; 1.  
DR PROSITE; PS60001; NOS; 1.  
FT NON TER 354 354  
SQ SEQUENCE 354 AA; 48824 MW; 8D76BAB87B739F9A CRC64;  
Query Match 100.0%; Score 35; DB 5; Length 354;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDTP 6

Db 109 GRGDTP 114

RESULT 2	
Q9AA55	
ID	Q9AA55 PRELIMINARY; PRT; 450 AA.
AC	Q9AA55
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 22, Last annotation update)
DE	Multidrug resistance protein, putative.
GN	CC0751
OS	Caulobacter crescentus.
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC	Caulobacteraceae; Caulobacter.
OX	NCBI_TaxID=155892;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=ATCC 19089 / CB15;
RC	MEDLINE=21173698; PubMed=11259647;
RX	Nierman W.C., Feidlyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA	Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.D., Haft D.H.,
RA	Kolchay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA	Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA	Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT	"Complete genome sequence of <i>Caulobacter crescentus</i> ."
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR	EMBL; AE005751; AAK22736.1; -
DR	PIR; D87342; D87342.
DR	TIGR; CC0751; -
DR	GO; GO:0016021; C: integral to membrane; IEA.
DR	GO; GO:0015520; F: tetracycline:hydrogen antiporter activity; IEA.
DR	GO; GO:005904; P: tetracycline transport; IEA.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR001958; TCR TetA.
DR	PRNTS; PR01035; TCRTEA.
DR	PROSITE; PS0850; MFS; 1.
KW	Complete proteome.
SQ	SEQUENCE 450 AA; 46965 MW; 66F7B879FB6F6892 CRC64;
Query Match 100.0%; Score 35; DB 16; Length 450;	
Best Local Similarity 100.0%; Pred. No. 78;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GRGDTTP 6
DB	220 GRGDTTP 225
RESULT 3	
Q9CSK5	
ID	Q9CSK5 PRELIMINARY; PRT; 490 AA.
AC	Q9CSK5
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	281000IM5Krik protein (Fragment).
GN	SF3B1 OR 281000IM05RIK.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6J; TISSUE=Embryo;
RC	MEDLINE=21085660; PubMed=11217851;
RX	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Query Match 100.0%; Score 35; DB 16; Length 496;	
Best Local Similarity 100.0%; Pred. No. 86;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GRGDTTP 6
DB	269 GRGDTTP 274
RESULT 5	
Q7ZTA4	
ID	Q7ZTA4 PRELIMINARY; PRT; 500 AA.
AC	Q7ZTA4
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Similar to splicing factor (Fragment).
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
[1]	
RN	SEQUENCE FROM N.A.
RP	TISSUE=Embryo;
RC	

```

RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC043874; AAH43874.1; -
FT NON TER 500 500
SQ SEQUENCE 500 AA; 54928 MW; 6D2B27FEB8C8FFAA CRC64;
Query Match 100.0%; Score 35; DB 13; Length 500;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRGDTP 6
DB 272 GRGDTP 277
|||||
RESULT 6
Q89P25 PRELIMINARY; PRT; 575 AA.
AC Q89P25;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B113658 protein.
OS B113658.
GN Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriuguchi X., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL: AF005948; BAC48923.1; -
DR GO: GO:0015036; Fdisulfide oxidoreductase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001327; FAD_Pyr_redox.
DR PRINTS: PR001100; Pyr_redox.
DR PRINTS: PR00368; FADPNR.
DR PRINTS: PR00411; ENDRDTASEI.
KW Complete proteome.
SQ SEQUENCE 575 AA; 60935 MW; 9785EBC554B26E92 CRC64;
Query Match 100.0%; Score 35; DB 16; Length 575;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRGDTP 6
DB 462 GRGDTP 467
|||||
RESULT 7
Q7Z497 PRELIMINARY; PRT; 789 AA.
AC Q7Z497;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;

```

```

EX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshitoki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.E.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC056155; AAH56155.1; -
KW Hypothetical protein.
FT NON TER 789 789
SQ SEQUENCE 789 AA; 87409 MW; 12193BA71838B32E CRC64;
Query Match 100.0%; Score 35; DB 4; Length 789;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRGDTP 6
DB 269 GRGDTP 274
|||||
RESULT 8
Q8T8C0 PRELIMINARY; PRT; 1209 AA.
AC Q8T8C0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitric oxide synthase.
GN BMNOS.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RA Imamura M., Yang J., Yamakawa M.;
RT "cDNA cloning, characterization and gene expression of nitric oxide
synthase from the silkworm, Bombyx mori.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB071182; BAB85836.1; -
DR GO: GO:0005489; F:electron transporter activity; IEA.
DR GO: GO:0004517; F:nitric-oxide synthase activity; IEA.
DR GO: GO:0016491; F:nitric-oxide synthase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0006809; P:nitric oxide biosynthesis; IEA.
DR InterPro: IPR003097; FAD_binding.
DR InterPro: IPR001094; Flavodoxin-like.
DR InterPro: IPR008254; Flav_nitox_synth.
DR InterPro: IPR001709; FPN_Cyt_redctse.
DR InterPro: IPR004030; NO_synthase.
DR InterPro: IPR001433; OxRed_FAD/NAD(P).
DR Pfam: PF00667; FAD_binding_1; 1.

```



```

RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AF003600; BAB78927.1; -.
DR PIR; AC2508; AC2508.
DR GO; GO:0046822; C:extrachromosomal DNA; IEA.
DR GO; GO:0005786; C:signal recognition particle; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0006605; P:protein targeting; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR008629; GUN4.
DR InterPro; IPR000897; SRP54.
DR Pfam; PF00023; ank; 10.
DR Pfam; PF05429; GUN4; 1.
DR SMART; SM00248; ANK; 9.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00300; SRP54; 1.
DR KMW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 606 AA; 66738 MW; 54A8102A5802B82 CRC64;

Query Match 91.4%; Score 32; DB 16; Length 606;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6
DB 120 GKGDTTP 125

RESULT 13
Q7VX39 PRELIMINARY; PRT; 749 AA.
AC Q7VX39
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative extracellular solute-binding protein.
GN Bp193.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tchana I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebatina M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.N., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skellon J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640417; CAE42263.1; -.
KW Complete proteome.
SQ SEQUENCE 749 AA; 84250 MW; F0F20CECB60DE9E CRC64;

Query Match 91.4%; Score 32; DB 16; Length 749;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GRGDTP 6
DB 432 GKGDTTP 437

RESULT 14
Q82BH0 PRELIMINARY; PRT; 1514 AA.
AC Q82BH0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN SAV5735.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMS 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMS 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinoe M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005044; BAC73447.1; -.
DR InterPro; IPR008979; Gal_bind_like.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1514 AA; 159051 MW; 05F6C676BEC428C7 CRC64;

Query Match 91.4%; Score 32; DB 16; Length 1514;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTP 6
DB 1509 GKGDTTP 1514

RESULT 15
Q7YU00 PRELIMINARY; PRT; 4245 AA.
AC Q7YU00
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dynein heavy chain, putative.
GN TB927.2.5270.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=569;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUTa10.1;
RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,
RA Biteau N., Khalak H.G., Lin X., Mason T., Hannick L., Caler E.,
RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pai G.,
RA Van Aken S., Uterback T., Haas B., Koo H.L., Umayam L., Suh B.,

```

RA Gerrard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,  
RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,  
RA Adams M.D., Fraser C.M., Donelson J.E.;  
RT "The sequence and analysis of Trypanosoma brucei chromosome II.";  
RL Nucleic Acids Res. 0:0-0(2003).  
DR EMBL; AE017170; AAQ15985.1; -;  
SQ SEQUENCE 4246 AA; 485740 MW; 5C2AF1D9ESC3177F CRC64;

Query Match 91.4%; Score 32; DB 5; Length 4246;  
Best Local Similarity 83.3%; Pred.No. 3.2e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDTTP 6  
|:|||||  
Db 428 GXGDTTP 433

Search completed: April 16, 2004, 07:27:53  
Job time : 48.4 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 55.3 Seconds  
(without alignments)  
39.939 Million cell updates/sec

Title: US-09-991-588B-10  
Perfect score: 43  
Sequence: 1 GRGDSPC 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 3  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 25: \*  
1: sp archaea: \*  
2: sp bacteria: \*  
3: sp fungi: \*  
4: sp human: \*  
5: sp invertebrate: \*  
6: sp mammal: \*  
7: sp mhc: \*  
8: sp organelle: \*  
9: sp phage: \*  
10: sp plant: \*  
11: sp rodent: \*  
12: sp virus: \*  
13: sp vertebrate: \*  
14: sp unclassified: \*  
15: sp virus: \*  
16: sp bacterioph: \*  
17: sp archaea: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	100.0	86	2	Q9E2D2
2	40	93.0	76	16	Q7ULR4
3	37	86.0	388	11	Q8BIZ9
4	37	86.0	704	10	Q8L7T0
5	37	86.0	704	10	Q9SUD8
6	37	86.0	999	15	Q83C50
7	37	86.0	1048	6	Q9XT27
8	35	81.4	572	16	Q89HB8
9	35	81.4	1126	5	Q9BPR0
10	35	81.4	1516	10	Q7XL85
11	35	81.4	1827	16	P96203
12	35	81.4	1827	16	Q7XLL7
13	34	79.1	23	4	Q9UC00
14	34	79.1	83	16	Q7WGA3
15	34	79.1	83	16	Q7WAS7
16	34	79.1	83	16	Q7WU91

17	34	79.1	189	9	Q8SCU8	Q8scu8 pseudomonas
18	34	79.1	208	16	Q92K28	Q92k28 rhizobium m
19	34	79.1	228	16	Q916L5	Q916l5 pseudomonas
20	34	79.1	268	15	Q8UTL4	Q8utl4 human immun
21	34	79.1	293	6	Q9XSG0	Q9xsg0 oryctolagus
22	34	79.1	350	16	Q7WJ62	Q7wj62 bordetella
23	34	79.1	350	16	Q7WA30	Q7wa30 bordetella
24	34	79.1	350	16	Q7VXD1	Q7vxd1 bordetella
25	34	79.1	363	11	Q8CCB5	Q8ccb5 mus musculu
26	34	79.1	376	15	Q8J4T2	Q8j4t2 human immun
27	34	79.1	432	15	Q994Q3	Q994q3 human immun
28	34	79.1	441	17	Q9HON2	Q9hqn2 halobacteri
29	34	79.1	478	10	Q7XLG7	Q7xlg7 oryza sativ
30	34	79.1	525	17	Q9HMF2	Q9hmf2 halobacteri
31	34	79.1	562	4	Q8NC49	Q8nc49 homo sapien
32	34	79.1	584	11	Q8K3K8	Q8k3k8 mus musculu
33	34	79.1	618	17	Q8TVQ5	Q8tyq5 methanopyru
34	34	79.1	659	10	Q8GWX7	Q8gwx7 arabidopsis
35	34	79.1	691	10	Q9M334	Q9m334 arabidopsis
36	34	79.1	979	15	Q901W9	Q901w9 human immun
37	34	79.1	997	15	Q8UT34	Q8ut34 human immun
38	34	79.1	999	15	Q994J9	Q994j9 human immun
39	34	79.1	999	15	Q8UT52	Q8ut52 human immun
40	34	79.1	1001	15	Q8UTR8	Q8utr8 human immun
41	34	79.1	1005	15	Q90CZ4	Q90c24 human immun
42	34	79.1	1005	15	Q8UTU5	Q8utu5 human immun
43	34	79.1	1009	15	Q8UT97	Q8ut97 human immun
44	34	79.1	1011	15	Q8UTP1	Q8utp1 human immun
45	34	79.1	1429	15	Q9WF84	Q9wf84 human immun

## ALIGNMENTS

## RESULT 1

Q9E2D2 PRELIMINARY; PRT; 86 AA.  
ID Q9E2D2  
AC Q9E2D2; 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RUI1/001.  
RA Muschler P.F.J., Endter C.;  
RT "Isolation of mcp genes from Sinorhizobium meliloti.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF238190; AAG37856.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 86 AA; 9128 MW; 87C5BF56B175DD2A CRC64;

Query Match 100.0%; Score 43; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. NO. 3.69;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7  
Db 35 GRGDSPC 41

## RESULT 2

Q7ULR4 PRELIMINARY; PRT; 76 AA.  
ID Q7ULR4  
AC Q7ULR4;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.

```

GN RB9342
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=1117;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=1.
RC MEDLINE=42735913; PubMed=12835415;
RX Gloeckner E.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RC "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1"; Acad. Sci. U.S.A. 100:8298-8303(2003).
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294149; CAD76205.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 76 AA; 8843 MW; BAB79AAF1D2BD765 CRC64;

Query Match 93.0%; Score 40; DB 16; Length 76;
Best Local Similarity 85.7%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 1 GRGDSPC 7
|:|||||
Db 14 GKGDSPC 20

RESULT 3
Q8BI29 PRELIMINARY; PRT; 388 AA.
AC Q8BI29
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CDNA FLJ10480 FIS.
GN CHD1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The RANOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RJ Nature 420:563-573(2002).
DR EMBL; AK048742; BAC33442.1; -.
DR MGD; MGI:1915308; Chd1l
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00487; DEXDC; 1.
SQ SEQUENCE 388 AA; 44478 MW; 97404639CFB1F5FF CRC64;

Query Match 85.0%; Score 37; DB 11; Length 388;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGSDSPC 7
|:|||||
Db 380 RGSDSPC 365

RESULT 4
Q8L7T0 PRELIMINARY; PRT; 704 AA.
ID Q8L7T0

```

```

AC Q8L7T0
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT4928010/T13J8.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Shinn P., Chen H., Cheuk R., Kim C.J., Banh C., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Tang C.C., Toriumi M., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY128275; AAM91084.1; -.
DR InterPro; IPR003016; Lipoyl_BS.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF01535; PPR; 17.
DR TIGRFAMs; TIGR00756; PPR; 16.
DR PROSITE; PS00189; LIPOYL; 1.
SQ SEQUENCE 704 AA; 79066 MW; C65F6D169D8B23A8 CRC64;

Query Match 86.0%; Score 37; DB 10; Length 704;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGSDSPC 7
|:|||||
Db 276 RGSDSPC 281

RESULT 5
Q9SUD8 PRELIMINARY; PRT; 704 AA.
ID Q9SUD8
AC Q9SUD8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T13J8.120 OR AT4928010.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Bevan M., Pohl T., Weizenegger T., Hoheisel J., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Pohl T., Weizenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035524; CAB36770.1; -.
DR EMBL; AL161572; CAB79603.1; -.
DR PIR; T02902; T02902.

```

DR InterPro: IPR003016; Lipcyl\_BS.  
 DR InterPro: IPR002885; PPR.  
 DR InterPro: IPR008941; TPR-like.  
 DR Pfam: PF01535; PPR; 17.  
 DR TIGRFAMs: TIGR00756; PPR; 16.  
 DR PROSITE: PS00189; LIPOYL; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 704 AA; 79109 MW; 9B5A7DB47647204E CRC64;

Query Match 86.0%; Score 37; DB 10; Length 704;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7  
 DB 276 RGDSPC 281

RESULT 6  
 Q8JC50 PRELIMINARY; PRT; 999 AA.  
 ID Q8JC50;  
 AC Q8JC50;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Pol protein (Fragment).  
 GN POL.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=00KE K2E2012;  
 RA Dowling W.E., Kim B., Mason C.J., Kasuma K.Monique., Alam U.,  
 RA Elson L., Birx D.L., Robb M.L., McCutchan F.E., Carr J.K.;  
 RT "Forty-one near full length HIV-1 sequences from Kenya reveal an  
 RT epidemic of subtype A and A-containing recombinants."  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBC databases.  
 CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS  
 CC [BY SIMILARITY].  
 DR EMBL: AF457054; AAN03043.1; -.  
 DR PIR: A35890; A35890.  
 DR GO: GO:0004190; F.aspartic-type endopeptidase activity; IEA.  
 DR GO: GO:0003677; F.DNA binding; IEA.  
 DR GO: GO:0004519; F.endonuclease activity; IEA.  
 DR GO: GO:0016787; F.hydrolase activity; IEA.  
 DR GO: GO:0008907; F.integrase activity; IEA.  
 DR GO: GO:0004523; F.ribonuclease H activity; IEA.  
 DR GO: GO:0003723; F.RNA binding; IEA.  
 DR GO: GO:0003964; F.RNA-directed DNA polymerase activity; IEA.  
 DR GO: GO:0016740; F.transferase activity; IEA.  
 DR GO: GO:0008270; F.zinc ion binding; IEA.  
 DR GO: GO:0015074; P.DNA integration; IEA.  
 DR GO: GO:0006310; P.DNA recombination; IEA.  
 DR GO: GO:0006278; P.RNA dependent DNA replication; IEA.  
 DR InterPro: IPR001969; Asparticase\_AS.  
 DR InterPro: IPR001037; Integrase\_C.  
 DR InterPro: IPR003308; Integrase\_Zn.  
 DR InterPro: IPR001995; Peptidase\_A2.  
 DR InterPro: IPR009007; Pept A acid.  
 DR InterPro: IPR002156; RNaseH.  
 DR InterPro: IPR001584; Rve.  
 DR InterPro: IPR000477; RVTse.  
 DR Pfam: PF00552; Integrase; 1.  
 DR Pfam: PF02022; Integrase\_Zn; 1.  
 DR Pfam: PF00075; rnaseH; 1.  
 DR Pfam: PF00665; rve; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP PROTEASE; 1.  
 DR PROSITE: PS00175; ASP PROT RETROV; 1.  
 KW Endonuclease; Hydrolase; Nucleotidyltransferase; Polypeptide; Polyprotein;

KW RNA-directed DNA polymerase; Transferase.  
 FT NON TER 1  
 SQ SEQUENCE 999 AA; 113218 MW; C1DF0E9D40D40AF4 CRC64;

Query Match 86.0%; Score 37; DB 15; Length 999;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7  
 DB 34 RGDSPC 39

RESULT 7  
 Q9XT27 PRELIMINARY; PRT; 1048 AA.  
 ID Q9XT27;  
 AC Q9XT27;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ceruloplasmin.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=99384006; PubMed=10452945;  
 RA Lockhart P.J., Mercer J.F.B.;  
 RT "Cloning and expression analysis of the sheep ceruloplasmin cDNA."  
 RL Gene 236:251-257(1999).  
 DR EMBL: AF134814; AAD41477.1; -.  
 DR HSBP; P00450; IKCW.  
 DR GO: GO:0005507; F.copper ion binding; IEA.  
 DR InterPro: IPR001117; Cu-oxidase.  
 DR InterPro: IPR008972; Cupredoxin.  
 DR InterPro: IPR002355; Cu ox copper\_BS.  
 DR Pfam: PF00394; Cu-oxidase; 3.  
 DR PROSITE: PS00079; MULTICOPPER OXIDASE1; 3.  
 DR PROSITE: PS00080; MULTICOPPER OXIDASE2; 1.  
 DR PROSITE: PS00080; MULTICOPPER OXIDASE2; 1.  
 SQ SEQUENCE 1048 AA; 119125 MW; 925F16D7B0549CBB CRC64;

Query Match 86.0%; Score 37; DB 6; Length 1048;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RGDSPC 7  
 DB 862 RGDSPC 868

RESULT 8  
 Q89HB8 PRELIMINARY; PRT; 572 AA.  
 ID Q89HB8;  
 AC Q89HB8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE BL16076 protein.  
 GN BL16076;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA 110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,

RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AP005957; BAC51341.1; -;  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR000873; AMP-bind.  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 572 AA; 61728 MW; 5BAEBCF974944063 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 572;  
 Best Local Similarity 71.4%; Pred. No. 1.6e-02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 DB 75 CHGDAFC 51  
 ID Q9BPR0 PRELIMINARY; PRT; 1126 AA.  
 AC Q9BPR0;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Receptor type guanylyl cyclase.  
 GN BM-GCY.  
 OS Bombyx mori (Silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
 OC Bombycidae; Bombyx.  
 CX NCBI\_TaxID=7091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shintarou T., Shigeki S., Taeko S., Yasuhisa E., Takaaki N.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21376156; PubMed=11483433;  
 RA Tanoue S., Sumida S., Suetesugu T., Endo Y., Nishioka T.;  
 RT "Identification of a receptor type guanylyl cyclase in the antennal  
 RT lobe and antennal sensory neurons of the silkworm, Bombyx mori.";  
 RL Insect Biochem. Mol. Biol. 31:971-973(2001).  
 DR HSSP; P26769; IAB8.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0005488; F:binding; IEA.  
 DR GO; GO:0004383; F:guanylate cyclase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0008528; F:peptide receptor activity, G-protein coupled; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0008075; F:receptor guanylate cyclase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR001828; ANF\_receptor.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR InterPro; IPR001993; Mitoch carrier.  
 DR InterPro; IPR001170; Ntpep receptorN.  
 DR InterPro; IPR000719; Prot Kinase.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF01094; ANF\_receptor; 1.  
 DR Pfam; PF00211; guanylate\_cyc; 2.  
 DR Pfam; PF00363; pkinase; 1.  
 DR PRINTS; PR00255; NATPEPTIDER.  
 DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00044; CYCC; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS01025; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE; PS00215; MITOCH\_CARRIER; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 SQ SEQUENCE 1126 AA; 126212 MW; F6989C5B0E0A9488 CRC64;

Query Match 81.4%; Score 35; DB 5; Length 1126;  
 Best Local Similarity 71.4%; Pred. No. 3.1e-02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 DB 197 GRGNSPC 203  
 ID Q7XL85 PRELIMINARY; PRT; 1516 AA.  
 AC Q7XL85;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE OSUNB0014D23.4 protein.  
 GN OSUNB0014D23.4.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Xiang H., Chen X.Y.,  
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,  
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,  
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,  
 RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,  
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
 RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,  
 RA Zhang R.Q., Guan J.P., Hong G.F.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL731642; CAB05270.1; -;  
 SQ SEQUENCE 1516 AA; 166386 MW; F48397FC3A65C16E CRC64;

Query Match 81.4%; Score 35; DB 10; Length 1516;  
 Best Local Similarity 71.4%; Pred. No. 4.1e-02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 DB 1009 GRGDRPC 1015  
 ID P96203 PRELIMINARY; PRT; 1827 AA.  
 AC P96203;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE PKSE (Polyketide synthase).  
 GN PPSP OR RV2934 OR MTCY19H9.02 OR MT3004.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 CX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.B. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Levin K., Felwell I., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Lages K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.",  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.,  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains",  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z83857; CAB06093.1; -;  
 DR EMBL; AE007122; AAK47331.1; -;  
 DR PIR; B70984; B70984.  
 DR HSSP; P25715; 1MLA.  
 DR TIGR; MT3004; -;  
 DR TubercuList; Rv2934; -;  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001227; Ac trans.  
 DR InterPro; IPR000794; Ketoacyl synth.  
 DR InterPro; IPR006162; Pp bind.  
 DR InterPro; IPR006163; Pp bind.  
 DR Pfam; PF00698; Acyl\_transf; 1.  
 DR Pfam; PF02801; ketoacyl-synt; 1.  
 DR Pfam; PF02801; ketoacyl-synt C; 1.  
 DR Pfam; PF00550; pp-binding; 1.  
 DR PROSITE; PS00075; ACP DOMAIN; 1.  
 DR PROSITE; PS00606; B-KETOACYL SYNTHASE; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
 KW Phosphopantetheine; Transferase; Complete proteome.  
 SQ SEQUENCE 1827 AA; 193312 MW; 88A3C1392838D953 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 1827;  
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 Db 366 GQGESPC 372

## RESULT 12

Q7TXL7

ID Q7TXL7 PRELIMINARY; PRT; 1827 AA.

AC Q7TXL7

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Phenolphthalein synthesis type-I polyketide synthase PPSD.

GN PPSD OR MB2959.

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1765;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AF2122/97.

RX MEDLINE=22769107; PubMed=12788972;

RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simcn S.,

RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,

RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RT "The complete genome sequence of Mycobacterium bovis.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 DR EMBL; BX248344; CAD96646.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 1827 AA; 193237 MW; 3081356D90C3FE71 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 1827;  
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7

Db 366 GQGESPC 372

## RESULT 13

Q9UC00

ID Q9UC00 PRELIMINARY; PRT; 23 AA.

AC Q9UC00

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Enhancement of wound HEALING process.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95130623; PubMed=7829572;

RA Pierschbacher M.D., Polarek J.W., Craig W.S., Tschopp J.P.,

RA Sipes N.J., Harper J.R.;

RL J. Cell. Biochem. 56:150-154 (1994).

DR GO; GO:0009611; P:response to wounding; TAS.

SQ SEQUENCE 23 AA; 2268 MW; CE73999CB903891 CRC64;

Query Match

Best Local Similarity 79.1%; Score 34; DB 4; Length 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 6

Db 1 GRGDSPC 6

## RESULT 14

Q7WGA3

ID Q7WGA3 PRELIMINARY; PRT; 83 AA.

AC Q7WGA3

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative cell division topological specificity factor.

GN MINE OR BB4016.

OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI\_TaxID=518;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RB50 / ATCC BAA-588;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Feitwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,

RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinovitch E., Rutter S., Saunders M., Saunders D., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,

RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 LR EMBL; BX640449; CAC34379.1; -  
 KW Cell division; Complete proteome.  
 SQ SEQUENCE 83 AA; 9489 MW; 98431A7A7BDD2F58 CRC64;

Query Match 79.1%; Score 34; DB 16; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 31 GRGDSP 36

## RESULT 15

Q7W4S7

ID Q7W4S7 PRELIMINARY; PRT; 83 AA.  
 AC Q7W4S7;  
 DT 01-OCT-2003 (TREMREL. 25, Created)  
 DT 01-OCT-2003 (TREMREL. 25, last sequence update)  
 DT 01-OCT-2003 (TREMREL. 25, last annotation update)  
 DE Putative cell division topological specificity factor.  
 GN MINE OR BP23581.  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=2910271;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bertley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin K., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640434; CAC38865.1; -  
 KW Cell division; Complete proteome.  
 SQ SEQUENCE 83 AA; 9489 MW; 98431A7A7BDD2F58 CRC64;

Query Match 79.1%; Score 34; DB 16; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 Db 31 GRGDSP 36

Search completed: April 16, 2004, 07:27:52  
 Job time : 57.3 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	34	79.1	285	1	YX47	STRCO	Q9x8h0 streptomyc
2	34	79.1	1256	1	FINC	CHICK	P11722 gallus gall
3	34	79.1	1328	1	FINC	ELEWA	Q91289 pleurodeles
4	34	79.1	2265	1	FINC	BOVIN	P07589 bos taurus
5	34	79.1	2386	1	FINC	HUMAN	P02751 homo sapien
6	34	79.1	2477	1	FINC	MOUSE	P11276 mus musculu
7	34	79.1	2477	1	FINC	RAT	P04937 rattus norv
8	34	79.1	2481	1	FINC	XENLA	Q91740 xenopus lae
9	33	76.7	277	1	YF99	MEIUSE	Q59994 methanococc
10	33	76.7	592	1	FZD9	MOUSE	Q9x216 mus musculu
11	32	74.4	82	1	TK32	PHONI	O76201 phoneutria
12	32	74.4	239	1	TKK2	CAVEO	P12323 cavia porce
13	32	74.4	270	1	FPG	PSEPK	Q88cq5 pseudomonas
14	32	74.4	270	1	FPG	PSEPK	Q88ah6 pseudomonas
15	32	74.4	337	1	MAT2	COCSA	Q9x445 cochl-obolu
16	32	74.4	340	1	MYB1	MAIZE	Q20024 zea mays (m
17	32	74.4	379	1	AUR	RAIZE	Q9x0r07 rhizobium m
18	32	74.4	391	1	AUR1	AGRT5	P58736 agrobacteri
19	32	74.4	427	1	IRKC	MOUSE	P52187 mus musculu
20	32	74.4	427	1	IRKC	RAT	P52188 rattus norv
21	32	74.4	432	1	IRKC	HUMAN	Q14500 homo sapien
22	32	74.4	432	1	IRK1	HUMAN	Q15756 homo sapien
23	32	74.4	474	1	UCP2	EUGGR	P43265 euglena gra
24	32	74.4	591	1	FZD9	HUMAN	O00144 homo sapien
25	32	74.4	831	1	SYFB	MYCTU	P94985 mycobacteri
26	32	74.4	1155	1	CIAB	BACTK	P06578 bacillus th
27	32	74.4	1155	1	N133	MOUSE	Q87099 mus musculu
28	32	74.4	1174	1	C1FA	BACTA	Q03746 bacillus th
29	32	74.4	1411	1	Y297	HUMAN	O15040 homo sapien
30	32	74.4	2715	1	MLL4	HUMAN	Q9umh6 homo sapien
31	31	72.1	303	1	CBPB	ASTFL	P04069 astacus flu
32	31	72.1	352	1	TRUD	PSEPK	Q88mf2 pseudomonas
33	31	72.1	352	1	TRUD	PSEPK	Q88f16 pseudomonas

FIMC CHICK  
 ID \_FIMC\_CHICK STANDARD; PRT; 1256 AA.  
 AC P1722; Q90921;  
 DT 01-CC1-1989 (Rel. 12, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Fibronection (FN) (Fragments).  
 GN FN.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 CC NCBI\_TaxID=9031;  
 [1]  
 RN SEQUENCE OF 1-50 FROM N.A.  
 RX MEDLINE=893117850; PubMed=6572007;  
 RA Hirano H., Yamada Y., Sullivan M., de Crombrughe B., Pastan I.,  
 RA Yamada K.M.;  
 RT "Isolation of genomic DNA clones spanning the entire fibronectin  
 RT gene";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:46-50(1983).  
 [2]  
 RN SEQUENCE OF 51-1256 FROM N.A.  
 RX STRAIN=White leghorn;  
 RA Norton P.A.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE OF 227-415 FROM N.A.  
 RX MEDLINE=96183658; PubMed=8603103;  
 RA Gehris A.L., Brandli D.W., Lewis S.D., Bennett V.D.;  
 RT "The exon encoding the fibronectin type III-9 repeat is  
 RT constitutively included in the mRNA from chick limb mesenchyme and  
 RT cartilage";  
 RL Biochim. Biophys. Acta 1311:5-12(1996).  
 [4]  
 RN SEQUENCE OF 327-599 FROM N.A.  
 RX MEDLINE=88050950; PubMed=2823899;  
 RA Kubomura S., Obara M., Karasaki Y., Taniguchi H., Gotoh S.,  
 RA Tsuda T., Higashi K., Ohsato K., Hiarno H.;  
 RT "Genetic analysis of the cell binding domain region of the chicken  
 RT fibronectin gene";  
 RL Biochim. Biophys. Acta 910:171-181(1987).  
 [5]  
 RN SEQUENCE OF 413-1256 FROM N.A.  
 RX MEDLINE=88142820; PubMed=2830487;  
 RA Norton P.A., Hynes R.O.;  
 RT "Alternative splicing of chicken fibronectin in embryos and in normal  
 RT and transformed cells";  
 RL Mol. Cell. Biol. 7:4297-4307(1987).  
 CC -1- FUNCTION: Fibronectins bind cell surfaces and various compounds  
 CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins  
 CC are involved in cell adhesion, cell motility, opsonization, wound  
 CC healing, and maintenance of cell shape.  
 CC -1- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED  
 CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;  
 CC TO A LESSER EXTENT HOMODIMERS.  
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=1;  
 CC Comment=A number of isoforms are produced. Each of the "extra  
 CC domain" and the connecting strand 3 are present in some forms of  
 CC fibronectin and absent in others;  
 CC Name=1;  
 CC IsoId=P11722-1; Sequence=displayed;  
 CC -1- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted  
 CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric  
 CC forms), made by fibroblasts, epithelial and other cell types, is  
 CC deposited as fibrils in the extracellular matrix.  
 CC -1- PTM: Sulfated (By similarity).  
 CC -1- SIMILARITY: Contains at least 2 fibronectin type I domains.  
 CC -1- SIMILARITY: Contains at least 8 fibronectin type III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V00432; CAA23714.1; -  
 DR EMBL; U21327; AAA73566.1; -  
 DR EMBL; X06533; CAA29781.1; -  
 DR EMBL; X26186; AAA48772.1; ALT\_SEQ.  
 DR EMBL; U23386; AAB01062.1; -  
 DR PIR; A28532; A28512.  
 DR PIR; A29355; A29355.  
 DR PIR; S71465; S71465.  
 DR HSSP; P02751; 1FNH.  
 DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR008957; FN III-like.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR InterPro; IPR003962; FNIII\_subd.  
 DR PRINTS; PR00014; FNTYPEIII.  
 DR SMART; SM00058; FN1; 3.  
 DR SMART; SM00060; FN3; 9.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; PARTIAL.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 2.  
 KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;  
 KW Repeat; Sulfation; Alternative splicing.  
 FT NON TER 1 51  
 FT DOMAIN 236 509  
 FT DOMAIN 690 961  
 FT DOMAIN 1153 1226  
 FT DOMAIN 327 415  
 FT DOMAIN 416 509  
 FT DOMAIN 510 599  
 FT DOMAIN 600 689  
 FT DOMAIN 690 781  
 FT DOMAIN 782 871  
 FT DOMAIN 872 961  
 FT DOMAIN 1083 1152  
 FT DOMAIN 1174 1218  
 FT DOMAIN 1219 >1256  
 FT SITE 493 495  
 FT DISULFID 1176 1205  
 FT DISULFID 1203 1215  
 FT DISULFID 1221 1248  
 FT CARBOHYD 122 122  
 FT CARBOHYD 1078 1078  
 FT CARBOHYD 1034 1034  
 FT CARBOHYD 1035 1035  
 FT CONFLICT 516 516  
 FT CONFLICT 569 572  
 FT NON TER 1256 1256  
 SQ SEQUENCE 1256 AA; 137435 MW; 345A4CA0E4D71D9B CRC64;  
 Query Match 79.1%; Score 34; DB 1; Length 1256;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GRGDSF 6  
 Db 492 GRGDSF 497  
 RESULT 3  
 ID \_FIMC\_PLEWA STANDARD; PRT; 1328 AA.  
 AC Q91289;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)



```

FT DISULFID 1268 1284 BY SIMILARITY.
FT DISULFID 1320 1320 INTERCHAIN (WITH C-1324) (BY SIMILARITY).
FT DISULFID 1324 1324 INTERCHAIN (WITH C-1320) (BY SIMILARITY).
FT SITE 461 463 CELL ATTACHMENT SITE.
FT CARBOHYD 89 89 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1328 AA; 145037 MW; E31BF7968A1D1E74 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 1328;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy 1 GRGDSP 6
Db 460 GRGDSP 465

RESULT 4
PINC_BOVIN
ID F1NC_BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT C1-APR-1988 (Rel. 07, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DB Fibronectin (FN).
DS FN1.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RK [1]
RP SEQUENCE.
RA MEDLINE=87054047; PubMed=3780752;
RX Skorstengaard K., Jensen M.S.; Sahl P., Petersen T.E., Magnusson S.;
RT "Complete primary structure of bovine plasma fibronectin.";
RL Eur. J. Biochem. 161:441-453(1986).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=83117805; PubMed=6218503;
RA Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,
RX Sahl P., Sottrup-Jensen L., Magnusson S.;
RT "Partial primary structure of bovine plasma fibronectin: three types
of internal homology.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
RN [3]
RP SEQUENCE OF 2170-2265 FROM N.A.
RX MEDLINE=83221567; PubMed=6304699;
RA Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;
RT "Isolation and characterization of cDNA clones for human and bovine
fibronectins.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
CC -!- FUNCTION: Fibronecting bind cell surfaces and various compounds
including collagen, fibrin, heparin, DNA, and actin. Fibronectins
are involved in cell adhesion, cell motility, opsonization, wound
healing, and maintenance of cell shape.
CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
TO A LESSER EXTENT HOMODIMERS.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=1;
Comment=A number of isoforms are produced. Each of the "extra
domain" and the connecting strand 3 are present in some forms of
fibronectin and absent in others;
CC Name=1;
CC isoID=P07589-1; Sequence=Displayed;
CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
forms), made by fibroblasts, epithelial and other cell types, is
deposited as fibrils in the extracellular matrix.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.

```



RN [2] SEQUENCE FROM N.A. (ISOFORMS 3; 7 AND 10).  
 RP TISSUE=Cervix;  
 RC Ansoorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
 RA Newes H.-W., Weil B., Amid C., Osanger A., Fobo G., Han M.,  
 RA Wiesmann S.;  
 RN Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3] SEQUENCE OF 1-38 FROM N.A.  
 RP MEDLINE=87030890; PubMed=3770189;  
 RX Gutman A., Yamada K.Y., Kornblith A.R.;  
 RA "Human fibronectin is synthesized as a pre-polypeptide.";  
 RL FEBS Lett. 207:145-148(1986).  
 RN [4] SEQUENCE OF 1-49 FROM N.A.  
 RP MEDLINE=87175578; PubMed=3031656;  
 RX Dean L.C., Bowles C.L., Bourgeois S.;  
 RA "Cloning and analysis of the promoter region of the human fibronectin  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:1876-1880(1987).  
 RN [5] SEQUENCE OF 28-2386 FROM N.A. (ISOFORM 3).  
 RP MEDLINE=85284965; PubMed=2992939;  
 RX Kornblith A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E.;  
 RA "Primary structure of human fibronectin: differential splicing may  
 generate at least 10 polypeptides from a single gene.";  
 RL EMBO J. 4:1755-1759(1985).  
 RN [6] SEQUENCE OF 103-481 AND 2228-2386 FROM N.A. (ISOFORMS 1; 3; 8 AND 9).  
 RP TISSUE=Peripheral blood T-cell, and Umbilical vein endothelial cells;  
 RC Godfrey H.P., Ebrahim A.A.;  
 RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7] SEQUENCE OF 973-2386 FROM N.A. (ISOFORM 3).  
 RP MEDLINE=84272258; PubMed=6462919;  
 RX Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;  
 RA "Human fibronectin: cell specific alternative mRNA splicing generates  
 polypeptide chains differing in the number of internal repeats.";  
 RL Nucleic Acids Res. 12:5853-5868(1984).  
 RN [8] SEQUENCE OF 1232-1782 FROM N.A. (ISOFORM 7).  
 RP MEDLINE=88233940; PubMed=3375063;  
 RX Paolella G., Henchcliffe C., Sebastio S., Baralle F.E.;  
 RA "Sequence analysis and in vivo expression show that alternative  
 splicing of ED-B and ED-A regions of the human fibronectin gene are  
 independent events.";  
 RL Nucleic Acids Res. 16:3545-3557(1988).  
 RN [9] SEQUENCE OF 1257-1365 FROM N.A. (ISOFORM 11).  
 RP MEDLINE=88041070; PubMed=3478690;  
 RX Gutman A., Kornblith A.R.;  
 RA "Identification of a third region of cell-specific alternative  
 splicing in human fibronectin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7179-7182(1987).  
 RN [10] SEQUENCE OF 1441-1548.  
 RP MEDLINE=82265604; PubMed=7050098;  
 RX Pierschbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;  
 RA "The cell attachment domain of fibronectin. Determination of the  
 primary structure.";  
 RL J. Biol. Chem. 257:9593-9597(1982).  
 RN [11] SEQUENCE OF 1448-1540 FROM N.A.  
 RP MEDLINE=83290929; PubMed=6688418;  
 RX Oldberg A., Linney E., Ruoslahti E.;  
 RA "Molecular cloning and nucleotide sequence of a cDNA clone coding for  
 the cell attachment domain in human fibronectin.";  
 RL J. Biol. Chem. 258:10193-10196(1983).  
 RN [12] SEQUENCE OF 1448-1540 FROM N.A.  
 RP MEDLINE=86111901; PubMed=3003095;  
 RX Oldberg A., Ruoslahti E.;  
 RA "Evolution of the fibronectin gene. Exon structure of cell attachment  
 domain.";  
 RL J. Biol. Chem. 261:2113-2116(1986).  
 RN [13] SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).  
 RP MEDLINE=85280409; PubMed=2992573;  
 RX Barnard M.P., Kolbe M., Weil D., Chu M.-L.;  
 RA "Human cellular fibronectin: comparison of the carboxyl-terminal  
 portion with rat identifies primary structural domains separated by  
 RT hypervariable regions.";  
 RL Biochemistry 24:2698-2704(1985).  
 RN [14] SEQUENCE OF 1712-1739 FROM N.A.  
 RP MEDLINE=87026578; PubMed=3021206;  
 RX Sekiguchi K., Kios A.M., Kurachi K., Yoshitake S., Hakomori S.;  
 RA "Human liver fibronectin complementary DNAs: identification of two  
 RT different messenger RNAs possibly encoding the alpha and beta  
 subunits of plasma fibronectin.";  
 RL Biochemistry 25:4936-4941(1986).  
 RN [15] SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).  
 RP TISSUE=Cartilage;  
 RC Parker A.E., Bontell J., Carr A., Maciewicz R.A.;  
 RA "Novel cartilage-specific splice variants of fibronectin.";  
 RT Osteoarthritis Cartilage 10:528-534(2002).  
 RN [16] SEQUENCE OF 32-290.  
 RP MEDLINE=84032463; PubMed=6630202;  
 RX Garcia-Pardo A., Pearlstein E., Frangione B.;  
 RA "Primary structure of human plasma fibronectin. The 29,000-dalton  
 RT NH2-terminal domain.";  
 RL J. Biol. Chem. 258:12670-12674(1983).  
 RN [17] SEQUENCE OF 309-608, AND COLLAGEN-BINDING.  
 RP MEDLINE=87080265; PubMed=3024962;  
 RX Owens R.J., Baralle F.E.;  
 RA "Mapping the collagen-binding site of human fibronectin by expression  
 RT in Escherichia coli.";  
 RL EMBO J. 5:2825-2830(1986).  
 RN [18] SEQUENCE OF 309-608, AND COLLAGEN-BINDING.  
 RP MEDLINE=86042625; PubMed=2414772;  
 RX Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
 RA "Tyrosine sulfation of proteins from the human hepatoma cell line  
 RT HepG2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
 RN [19] O-GLYCOSYLATION OF THR-2064.  
 RP MEDLINE=91190085; PubMed=2012601;  
 RX Tresselt T., McCarthy J.B., Calaycay J., Lee T.D., Jegesse K.,  
 RA Shively J.B., Pande H.;  
 RA "Human plasma fibronectin. Demonstration of structural differences  
 RT between the A- and B-chains in the III CS region.";  
 RL Biochem. J. 274:731-738(1991).  
 RN [20] FBLN1-BINDING SITE.  
 RP MEDLINE=93015879; PubMed=1400330;  
 RX Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,  
 RA Argraves W.S.;  
 RA "Fibulin binds to itself and to the carboxyl-terminal heparin-binding  
 RT region of fibronectin.";  
 RL J. Biol. Chem. 267:20120-20125(1992).  
 RN [21] CHARACTERIZATION OF FIBRIN-BINDING SITE 1.  
 RP MEDLINE=95081153; PubMed=7989369;  
 RX Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;  
 RA "Further characterization of the NH2-terminal fibrin-binding site on  
 RT fibronectin.";  
 RL J. Biol. Chem. 269:31938-31945(1994).  
 RN [22] STRUCTURE BY NMR OF 1447-1540.  
 RP MEDLINE=92162710; PubMed=1311202;  
 RX Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,  
 RA

RA Campbell I.D.;  
RT "1H NMR assignment and secondary structure of the cell adhesion type  
RT III module of fibronectin.";  
RL Biochemistry 31:2068-2073(1992).  
[23]  
RP STRUCTURE BY NMR OF 1447-1540.  
RX MEDLINE=93046665; PubMed=1423622;  
RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;  
RT "The three-dimensional structure of the tenth type III module of  
RT fibronectin: an insight into RGD-mediated interactions.";  
RL Cell 71:671-678(1992).  
[24]  
RP STRUCTURE BY NMR OF 182-275.  
RX MEDLINE=94141923; PubMed=8038892;  
RA Williams M.J., Phan I., Harvey T.S., Roszagno A., Gold L.I.,  
RA Campbell I.D.;  
RT "Solution structure of a pair of fibronectin type I modules with  
RT fibrin binding activity.";  
RL J. Mol. Biol. 235:1302-1311(1994).  
[25]  
RP STRUCTURE BY NMR OF 32-92.  
RX MEDLINE=96069779; PubMed=7583666;  
RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;  
RT "High-resolution structural studies of the factor XIIIa crosslinking  
RT site and the first type I module of fibronectin.";  
RL Nat. Struct. Biol. 2:946-950(1995).  
[26]  
RP STRUCTURE BY NMR OF 406-454.  
RX MEDLINE=98179558; PubMed=9514732;  
RA Sticht H., Pickford A.R., Potts J.R., Campbell I.D.;  
RT "Solution structure of the glycosylated second type 2 module of  
RT fibronectin.";  
RL J. Mol. Biol. 276:177-187(1998).  
[27]  
RP STRUCTURE BY NMR OF EXTRA ED-B DOMAIN FROM ISOFORM 7.  
Query Match 79.1%; Score 34; DB 1; Length 2386;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GRGDSF 6  
Db 1523 GRGDSF 1528  
RESULT 6  
FIRC MOUSE  
ID FIRC MOUSE STANDARD; PRT; 2477 AA.  
AC P11276; Q61567; Q61568; Q61569; Q64233; Q80U14;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Fibronectin precursor (FN).  
GN FN1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE OF 1-920 FROM N.A.  
RX STRAIN=FVB/N-3; TISSUE=Breast tumor;  
RX MEDLINE=22388257; PubMed=12477912;  
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Aitschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heisler F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek T.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Scallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RP SEQUENCE OF 1-28 FROM N.A.  
RX TISSUE=Liver;  
RX MEDLINE=94131313; PubMed=8299972;  
RA Polly P., Nicholson R.C.;  
RT "Sequence of the mouse fibronectin-encoding gene promoter region.";  
RL Gene 137:353-354(1993).  
[3]  
RP SEQUENCE OF 562-834 FROM N.A.  
RX STRAIN=NMRI;  
RX MEDLINE=95403556; PubMed=7673336;  
RA Tait J.F., Weiler A., Timpl R., Ekblom M., Ekblom P.;  
RT "Regulation of mesenchymal extracellular matrix protein synthesis by  
RT transforming growth factor-beta and glucocorticoids in tumor  
RT stroma.";  
RL J. Cell Sci. 108:2153-2162(1995).  
[4]  
RP SEQUENCE OF 899-2376 FROM N.A.  
RX Gorski G., Aros M., Norton P.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE OF 2375-2477 FROM N.A.  
RX MEDLINE=88124987; PubMed=3124113;  
RA Elatti S.P., Foster D.N., Ranganathan G., Moses H.L., Getz X.J.;  
RT "Induction of fibronectin gene transcription and mRNA is a primary  
RT response to growth-factor stimulation of AKR-2B cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:1119-1123(1988).  
[6]  
RP SEQUENCE OF 2375-2477 FROM N.A.  
RX TISSUE=Kidney;  
RX MEDLINE=93011722; PubMed=1327855;  
RA Khandjian E.W., Salomon C., Leonard N., Tremblay S., Turler H.;  
RT "Fibronectin gene expression in proliferating, quiescent, and SV40-  
RT infected mouse kidney cells.";  
RL Exp. Cell Res. 202:464-470(1992).  
[7]  
RP STRUCTURE BY NMR OF 1447-1630.  
RX MEDLINE=98202578; PubMed=9533887;  
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,  
RA Pastor R.W., Krueger S., Torchia D.A.;  
RT "Solution structure and dynamics of linked cell attachment modules of  
RT mouse fibronectin containing the RGD and synergy regions: comparison  
RT with the human fibronectin crystal structure.";  
RL J. Mol. Biol. 277:663-682(1998).  
[8]  
RP DOWN-REGULATION BY GLUCOCORTICOID.  
RX MEDLINE=21600963; PubMed=11737251;  
RA Gu Y.-C., Tait J.F., Gullberg D., Timpl R., Ekblom M.;  
RT "Glucocorticoids down-regulate the extracellular matrix proteins  
RT fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";  
RL Eur. J. Haematol. 67:176-184(2001).  
CC -1- FUNCTION: Fibronectins bind cell surfaces and various compounds  
CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins  
CC are involved in cell adhesion, cell motility, opsonization, wound  
CC healing, and maintenance of cell shape.  
CC -1- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced  
CC variants, connected by 2 disulfide bonds near the carboxyl ends;  
CC to a lesser extent homodimers. Interacts with FBLN1 (By  
CC similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=1;  
CC Comment=A number of isoforms are produced. Each of the "extra  
CC domain" and the connecting strand 3 are present in some forms of

fibronection and absent in others;

Name=1;  
 isoId=p11276-1; Sequence=Displayed;  
 TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted by hepatocytes. Cellular FN (dimeric or cross-linked multimeric forms), made by fibroblasts, epithelial and other cell types, is deposited as fibrils in the extracellular matrix.

INDUCTION: Glucocorticoids suppressed mRNA expression and protein synthesis.

PTM: Sulfated (By similarity).

SIMILARITY: Contains 12 fibronection type I domains.

SIMILARITY: Contains 2 fibronection type II domains.

SIMILARITY: Contains 17 fibronection type III domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; BC051082; AAH51082.1; -  
 EMBL; Z22729; CAA80422.1; -  
 EMBL; X62452; CAA57796.1; -  
 EMBL; X93157; CAA63654.1; -  
 EMBL; M18134; AAB37636.1; -  
 EMBL; S45680; AAB23491.1; -  
 PIR; A49173; A49173.  
 PIR; I48349; I48349.  
 PDB; 1MFN; 29-APR-98.  
 PDB; 2MEN; 29-APR-98.  
 MGD; MGI:95566; Fn1.  
 GO; GO:0007155; P:cell adhesion; IDA.  
 InterPro; IPR006209; EGF like.  
 InterPro; IPR000083; Fibrnctnl.  
 InterPro; IPR008957; FN III-like.  
 InterPro; IPR003961; FN\_III.  
 InterPro; IPR000562; FN\_Type\_II.  
 InterPro; IPR003962; FnIII\_subd.  
 InterPro; IPR008924; MCR\_alpha\_beta\_C.  
 PRINTS; PRO0014; FNTYPEII.  
 SMART; SM00058; FN1; 4.  
 SMART; SM00060; FN3; 12.  
 PROSITE; PS00022; EGF\_1; 2.  
 PROSITE; PS00023; FIBRONECTIN\_2; 2.  
 PROSITE; PS01253; FIBRONECTIN\_1; 12.  
 Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation; Sulfation; Cell adhesion; Repeat; Alternative splicing; Signal; 3D-structure.

SIGNAL 1 32 BY SIMILARITY.  
 CHAIN 33 2477 FIBRONECTIN.  
 DOMAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.  
 DOMAIN 308 658 FIBRIN- AND HEPARIN-BINDING.  
 DNA\_BIND 906 1171  
 DOMAIN 1357 1630 CELL-ATTACHMENT.  
 DOMAIN 1811 2331 HEPARIN-BINDING 2.  
 DOMAIN 2296 2427 FIBRIN-BINDING 2.  
 DOMAIN 51 96 FIBRONECTIN TYPE-I 1.  
 DOMAIN 96 140 FIBRONECTIN TYPE-I 2.  
 DOMAIN 140 185 FIBRONECTIN TYPE-I 3.  
 DOMAIN 185 230 FIBRONECTIN TYPE-I 4.  
 DOMAIN 230 272 FIBRONECTIN TYPE-I 5.  
 DOMAIN 306 343 FIBRONECTIN TYPE-I 6.  
 DOMAIN 345 404 FIBRONECTIN TYPE-II 1.  
 DOMAIN 405 469 FIBRONECTIN TYPE-II 2.  
 DOMAIN 468 516 FIBRONECTIN TYPE-I 7.  
 DOMAIN 516 559 FIBRONECTIN TYPE-I 8.  
 DOMAIN 559 602 FIBRONECTIN TYPE-I 9.  
 DOMAIN 609 706 FIBRONECTIN TYPE-III 1.  
 DOMAIN 707 808 FIBRONECTIN TYPE-III 2.  
 DOMAIN 809 903 FIBRONECTIN TYPE-III 3.  
 DOMAIN 904 994 FIBRONECTIN TYPE-III 4.

FT	DOMAIN	995	1084	FIBRONECTIN TYPE-III 5.
FT	DOMAIN	1085	1172	FIBRONECTIN TYPE-III 6.
FT	DOMAIN	1173	1264	FIBRONECTIN TYPE-III 7.
FT	DOMAIN	1265	1355	FIBRONECTIN TYPE-III 8 (EXTRA DOMAIN 1).
FT	DOMAIN	1356	1446	FIBRONECTIN TYPE-III 9.
FT	DOMAIN	1447	1536	FIBRONECTIN TYPE-III 10.
FT	DOMAIN	1537	1630	FIBRONECTIN TYPE-III 11.
FT	DOMAIN	1631	1720	FIBRONECTIN TYPE-III 12.
FT	DOMAIN	1721	1810	FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN 2).
FT	DOMAIN	1811	1902	FIBRONECTIN TYPE-III 14.
FT	DOMAIN	1903	1991	FIBRONECTIN TYPE-III 15.
FT	DOMAIN	1992	2081	FIBRONECTIN TYPE-III 16.
FT	DOMAIN	2082	2201	CONNECTING STRAND 3 (CS-3) (V REGION).
FT	DOMAIN	2202	2283	FIBRONECTIN TYPE-III 17.
FT	DOMAIN	2294	2338	FIBRONECTIN TYPE-I 10.
FT	DOMAIN	2339	2381	FIBRONECTIN TYPE-I 11.
FT	DOMAIN	2383	2426	FIBRONECTIN TYPE-I 12.
FT	SITE	1614	1616	CELL ATTACHMENT SITE.
FT	SITE	2181	2183	CELL ATTACHMENT SITE.
FT	DISULFID	53	79	BY SIMILARITY.
FT	DISULFID	77	88	BY SIMILARITY.
FT	DISULFID	98	126	BY SIMILARITY.
FT	DISULFID	124	136	BY SIMILARITY.
FT	DISULFID	142	170	BY SIMILARITY.
FT	DISULFID	168	180	BY SIMILARITY.
FT	DISULFID	187	216	BY SIMILARITY.
FT	DISULFID	214	226	BY SIMILARITY.
FT	DISULFID	232	261	BY SIMILARITY.

Query Match 79.1%; Score 34; DB 1; Length 2477;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 GRGDSP 6

Db 1613 GRGDSP 1618

RESULT 7

FINC\_RAT

ID FINC\_RAT STANDARD; PRT; 2477 AA.

AC P04937;

DT 13-AUG-1987 (Rel. 05, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Fibronection precursor [FN].

GN FN1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=101116;

RN [1]

RC SEQUENCE FROM N.A.

RP STRAIN=Fischer; TISSUE=Liver;

RX MEDLINE=88054951; PubMed=2445560;

RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;

RT "Multiple sites of alternative splicing of the rat fibronection gene transcript.";

RL EMBL J. 6:2573-2580 (1987).

RN [2]

RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.

RC STRAIN=Fischer; TISSUE=Liver;

RX MEDLINE=88054950; PubMed=3119323;

RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;

RT "Organization of the fibronection gene provides evidence for exon shuffling during evolution.";

RL EMBL J. 6:2565-2572 (1987).

RN [3]

RP SEQUENCE OF 1586-2477 FROM N.A.

RX MEDLINE=84082067; PubMed=6317187;

RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;

RT "Three different fibronection mRNAs arise by alternative splicing within the coding region.";

FT	CHAIN	33	2477	FIBRONECTIN.
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23
24	24	24	24	24
25	25	25	25	25
26	26	26	26	26
27	27	27	27	27
28	28	28	28	28
29	29	29	29	29
30	30	30	30	30
31	31	31	31	31
32	32	32	32	32
33	33	33	33	33
34	34	34	34	34
35	35	35	35	35
36	36	36	36	36
37	37	37	37	37
38	38	38	38	38
39	39	39	39	39
40	40	40	40	40
41	41	41	41	41
42	42	42	42	42
43	43	43	43	43
44	44	44	44	44
45	45	45	45	45
46	46	46	46	46
47	47	47	47	47
48	48	48	48	48
49	49	49	49	49
50	50	50	50	50
51	51	51	51	51
52	52	52	52	52
53	53	53	53	53
54	54	54	54	54
55	55	55	55	55
56	56	56	56	56
57	57	57	57	57
58	58	58	58	58
59	59	59	59	59
60	60	60	60	60
61	61	61	61	61
62	62	62	62	62
63	63	63	63	63
64	64	64	64	64
65	65	65	65	65
66	66	66	66	66
67	67	67	67	67
68	68	68	68	68
69	69	69	69	69
70	70	70	70	70
71	71	71	71	71
72	72	72	72	72
73	73	73	73	73
74	74	74	74	74
75	75	75	75	75
76	76	76	76	76
77	77	77	77	77
78	78	78	78	78
79	79	79	79	79
80	80	80	80	80
81	81	81	81	81
82	82	82	82	82
83	83	83	83	83
84	84	84	84	84
85	85	85	85	85
86	86	86	86	86
87	87	87	87	

```

FT MOD_RES 2475 2475 PHOSPHORYLATION (BY SIMILARITY).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 876 876 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1006 1006 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 2198 2198 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 2154 2154 O-LINKED (GALNAC. . .) (BY SIMILARITY).
FT VARSPLIC 1720 1809 Missing (in isoform 2).
FT VARSPLIC 2082 2106 Missing (in isoform 3).
FT VARSPLIC 2082 2200 Missing (in isoform 4).
FT CONFLICT 2318 2318 G -> A (IN REF. 3).
SQ SEQUENCE 2477 AA; 272510 MW; B4391A472CEDEB5 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 2477;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSF 6
Db 1613 GRGDSF 1618

RESULT 8
PINC_XENLA STANDARD; PRT; 2481 AA.
AC Q91740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibronectin precursor.
GN FNI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=92111942; PubMed=1730390;
RA Desimone D.W., Norton P.A., Hynes R.O.;
RT "Identification and characterization of alternatively spliced
  fibronectin mRNAs expressed in early Xenopus embryos.";
RL Dev. Biol. 149:357-369(1992).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
  including collagen, fibrin, heparin, DNA, and actin. Fibronectins
  are involved in cell adhesion, cell motility, opsonization, wound
  healing, and maintenance of cell shape (By similarity).
CC -!- SUBUNIT: DIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS,
  CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS (BY
  SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS:
  Event=Alternative splicing; Named isoforms=1;
  Comment=A number of isoforms are produced. Each of the "extra
  domain" and the connecting strand 3 are present in some forms of
  fibronectin and absent in others;
  Name=1;
CC -!- TISSUE SPECIFICITY: IN EARLY XENOPUS EMBRYO, CELLULAR FORMS OF
  FIBRONECTIN PREDOMINATE WHICH INCLUDE BOTH EXTRA DOMAINS. IN
  FIBRONECTIN OF EMBRYONIC AND ADULT LIVER THE CONNECTING STRAND 3
  CAN BE EITHER COMPLETELY EXCLUDED OR INCLUDED.
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M77820; AAA49707.1; -.
CC HSSP; P02751; 2FN3.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000083; Fibnctnl.
CC InterPro; IPR008957; FN III-like.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR000562; FN_Type_II.
CC Pfam; PF00039; fn1; 12.
CC Pfam; PF00040; fn2; 2.
CC Pfam; PF00041; fn3; 17.
CC PRINTS; PR00013; FNTYPEII.
CC ProDom; PD000995; FN_Type_II; 2.
CC SMART; SM00058; FNI; 12.
CC SMART; SM00059; FN2; 2.
CC SMART; SM00060; FN3; 17.
CC PROSITE; PS00022; EGF 1; 2.
CC PROSITE; PS01253; FIBRONECTIN 1; 11.
CC Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;
  Repeat; Alternative splicing; Signal.
KW SIGNAL 1 31
FT CHAIN 32 2481 FIBRONECTIN.
FT DOMAIN 55 275 FIBRIN- AND HEPARIN-BINDING 1.
FT DOMAIN 309 609 COLLAGEN-BINDING.
FT DNA_BIND 907 1172 BY SIMILARITY.
FT DOMAIN 1358 1631 CELL-ATTACHMENT.
FT DOMAIN 1812 2082 HEPARIN-BINDING 2.
FT DOMAIN 2301 2432 FIBRIN-BINDING 2.
FT DOMAIN 53 93 FIBRONECTIN TYPE-I 1.
FT DOMAIN 98 141 FIBRONECTIN TYPE-I 2.
FT DOMAIN 142 185 FIBRONECTIN TYPE-I 3.
FT DOMAIN 187 231 FIBRONECTIN TYPE-I 4.
FT DOMAIN 232 276 FIBRONECTIN TYPE-I 5.
FT DOMAIN 307 346 FIBRONECTIN TYPE-I 6.
FT DOMAIN 346 405 FIBRONECTIN TYPE-II 1.
FT DOMAIN 406 470 FIBRONECTIN TYPE-II 2.
FT DOMAIN 469 512 FIBRONECTIN TYPE-I 7.
FT DOMAIN 517 559 FIBRONECTIN TYPE-I 8.
FT DOMAIN 560 603 FIBRONECTIN TYPE-I 9.
FT DOMAIN 610 707 FIBRONECTIN TYPE-III 1.
FT DOMAIN 708 809 FIBRONECTIN TYPE-III 2.
FT DOMAIN 810 904 FIBRONECTIN TYPE-III 3.
FT DOMAIN 905 995 FIBRONECTIN TYPE-III 4.
FT DOMAIN 996 1085 FIBRONECTIN TYPE-III 5.
FT DOMAIN 1086 1173 FIBRONECTIN TYPE-III 6.
FT DOMAIN 1174 1265 FIBRONECTIN TYPE-III 7.
FT DOMAIN 1266 1356 FIBRONECTIN TYPE-III 8 (EXTRA DOMAIN).
FT DOMAIN 1357 1447 FIBRONECTIN TYPE-III 9.
FT DOMAIN 1448 1537 FIBRONECTIN TYPE-III 10.
FT DOMAIN 1538 1631 FIBRONECTIN TYPE-III 11.
FT DOMAIN 1632 1721 FIBRONECTIN TYPE-III 12.
FT DOMAIN 1722 1811 FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN).
FT DOMAIN 1812 1903 FIBRONECTIN TYPE-III 14.
FT DOMAIN 1904 1992 FIBRONECTIN TYPE-III 15.
FT DOMAIN 1993 2082 FIBRONECTIN TYPE-III 16.
FT DOMAIN 2083 2205 CONNECTING STRAND 3 (CS-3) (V REGION).
FT DOMAIN 2206 2287 FIBRONECTIN TYPE-III 17.
FT DOMAIN 2299 2343 FIBRONECTIN TYPE-I 10.
FT DOMAIN 2344 2386 FIBRONECTIN TYPE-I 11.
FT DOMAIN 2388 2431 FIBRONECTIN TYPE-I 12.
FT SITE 1615 1617 CELL ATTACHMENT SITE.
FT DISULFID 55 81 BY SIMILARITY.
FT DISULFID 79 90 BY SIMILARITY.
FT DISULFID 160 128 BY SIMILARITY.
FT DISULFID 126 138 BY SIMILARITY.
FT DISULFID 144 172 BY SIMILARITY.
FT DISULFID 170 182 BY SIMILARITY.
FT DISULFID 189 218 BY SIMILARITY.
FT DISULFID 216 228 BY SIMILARITY.

```

FT DISULFID 234 263 BY SIMILARITY.  
 FT DISULFID 261 273 BY SIMILARITY.  
 FT DISULFID 309 336 BY SIMILARITY.  
 FT DISULFID 334 343 BY SIMILARITY.  
 FT DISULFID 361 387 BY SIMILARITY.  
 FT DISULFID 375 402 BY SIMILARITY.  
 FT DISULFID 421 447 BY SIMILARITY.  
 FT DISULFID 471 499 BY SIMILARITY.  
 FT DISULFID 497 509 BY SIMILARITY.  
 FT DISULFID 519 546 BY SIMILARITY.  
 FT DISULFID 544 556 BY SIMILARITY.  
 FT DISULFID 562 590 BY SIMILARITY.  
 FT DISULFID 588 600 BY SIMILARITY.  
 FT DISULFID 2301 2330 BY SIMILARITY.  
 FT DISULFID 2328 2340 BY SIMILARITY.  
 FT DISULFID 2346 2373 BY SIMILARITY.  
 FT DISULFID 2371 2383 BY SIMILARITY.  
 FT DISULFID 2390 2414 BY SIMILARITY.  
 FT DISULFID 2412 2428 BY SIMILARITY.  
 FT DISULFID 2459 2459 INTERCHAIN (WITH C-2463) (BY SIMILARITY).  
 FT DISULFID 2463 2463 INTERCHAIN (WITH C-2459) (BY SIMILARITY).  
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 877 877 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1244 1244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1291 1291 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2202 2202 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 2481 AA; 272678 MW; 7847DF4F6CE72C93 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 2481;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSP 6  
 |||||  
 1614 GRGDSP 1619

DB 1614 GRGDSP 1619

RESULT 9

YF99 METJA STANDARD; PRT; 277 AA.  
 AC Q58934;  
 DT 10-OCT-2003 (Rel. 42, Created;  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein MJ1599.  
 GN MJ1599.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL Science 273:1058-1073 (1996).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; U67600; AAB99626.1; -.  
 DR PIR; P64499; P64499.  
 DR TIGR; MJ1599; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 277 AA; 32340 MW; F59EF4A2485D3BAE CRC64;

Query Match 76.7%; Score 33; DB 1; Length 277;  
 Best Local Similarity 83.3%; Pred. No. 21;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7  
 |||||  
 131 RGDSPC 136

DB 131 RGDSPC 136

RESULT 10

FZD9 MOUSE STANDARD; PRT; 592 AA.  
 ID FZD9\_MOUSE  
 AC Q9R216; Q35494; Q9CX16; Q9R2B3;  
 DT 28-FEB-2003 (Rel. 41, Created;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Frizzled 9 precursor (Frizzled-3) (Fz-9) (mFz3).  
 GN FZD9 OR FZD3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.  
 RC STRAIN=129/SvJ; TISSUE=Brain;  
 RX MEDLINE=99216417; PubMed=10198163;  
 RA Wang Y.-K., Spoerle R., Paperna T., Schughart K., Francke U.;  
 RT "Characterization and expression pattern of the frizzled Gene Fzd9,  
 RT the mouse homolog of FZD9 which is deleted in Williams-Beuren  
 RT syndrome.";  
 RL Genomics 57:235-248 (1999).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryonic head;  
 RX MEDLINE=21085680; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Havaheizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 403:685-690 (2001).  
 [3]  
 RN SEQUENCE OF 44-592 FROM N.A.  
 RC STRAIN=JUN-1999; to the EMBL/GenBank/DBJ databases.  
 RA Van Raay T.J., Rasmussen J.T., Rao M.S.;  
 RT "A novel mouse frizzled gene expressed in early neural development.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN SEQUENCE OF 144-592 FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Calo L., Mimmack M.L., Keverne E.B., Emson P.C.;



RT "Localization of the mouse frizzled gene, mFZD3 in the olfactory epithelium and in the vomeronasal organ.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors are coupled to the beta-catenin canonical signaling pathway, which leads to the activation of dishevelled proteins, inhibition of GSK-3 kinase, nuclear accumulation of beta-catenin and activation of Wnt target genes. A second signaling pathway involving PKC and calcium fluxes has been seen for some family members, but it is not yet clear if it represents a distinct pathway or if it can be integrated in the canonical pathway, as PKC seems to be required for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem to involve interactions with G-proteins. May be involved in transduction and intercellular transmission of polarity information during tissue morphogenesis and/or in differentiated tissues.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes in the plasma membrane.

CC -!- TISSUE SPECIFICITY: In the embryo, found in the neural tube, trunk skeletal muscle precursors (myotomes), limb skeletal anlagen, craniofacial regions and nephric ducts. In the adult, expression is abundant in heart, brain, testis and skeletal muscle. In the testis, expressed in all spermatogenic cell types. Lower levels in adult lung, liver and kidney. Barely detectable in spleen.

CC -!- EXPRESSION: Expressed also in chondrocytes.

CC -!- DEVELOPMENTAL STAGE: Not detected at embryonic day 7 (E7), weakly at E11 and strongly at E15 and E17. Expression covers the entire neural tube at 9.5 dpc, decreases at 10.5 dpc and becomes detectable only in the lumbar to tail regions at 11.5 dpc. In the somites, expression begins at 10.5 dpc to become upregulated all along the rostrocaudal trunk axis at 11.5 dpc. In craniofacial territories, expression is first detected at 11.5 dpc in restricted areas of the nose, the maxillary mandibular and second branchial arch anlagen. At 11.5 dpc, predominantly expressed in restricted areas of the nose, dorsally to the eye and in the caudal pharyngeal region.

CC -!- DOMAIN: Lys-Thr-X-X-Trp motif is involved in the activation of the Wnt/beta-catenin signaling pathway (By similarity).

CC -!- DOMAIN: The FZ domain is involved in binding with Wnt ligands (By similarity).

CC -!- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor family.

CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.

CC -!- CAUTION: Has been first described as FZD3 in literature.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF088850; AAD27789.1; -

DR EMBL: AK021164; BAB3231.1; -

DR EMBL: AP033585; XAB87503.2; -

DR EMBL: Y17709; CAB44237.1; -

DR MGD: MGI:1313278; Fzd9.

DR InterPro: IPR000539; Frizzled.

DR InterPro: IPR000024; Fz domain.

DR InterPro: IPR000832; GPCR secretin.

DR Pfam: PF01534; Frizzled; 1.

DR Pfam: PF01392; Fz; 1.

DR PRINTS: PR00489; FRIZZLED.

DR SMART: SM00063; FRI; 1.

DR PROSITE: PS00318; FZ; 1.

DR PROSITE: PS0261; G-PROTEIN\_RECEPTOR\_FZ\_4; 1.

KW Multigene family; G-protein coupled receptor; Transmembrane; Signal.

KW Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.

FT SIGNAL 1 23

FT CHAIN 24 592

FT DOMAIN 24 230

FT TRANSMEM 231 251

FT DOMAIN 252 267

FT TRANSMEM 268 288

FT DOMAIN 289 316

FT TRANSMEM 317 337

FT DOMAIN 338 356

FT TRANSMEM 357 377

FT DOMAIN 378 401

FT TRANSMEM 402 422

FT DOMAIN 423 448

FT TRANSMEM 449 469

FT DOMAIN 470 509

FT TRANSMEM 510 530

FT DOMAIN 531 592

FT SITE 35 156

FT SITE 533 538

FT CARBOHYD 54 54

FT CARBOHYD 159 159

FT CONFLICT 66 66

FT CONFLICT 73 74

FT CONFLICT 93 93

FT CONFLICT 144 144

FT CONFLICT 221 221

FT CONFLICT 237 237

FT CONFLICT 308 308

FT CONFLICT 374 374

FT CONFLICT 592 592

SQ SEQUENCE 592 AA; 64994 MW; 21B2D4F8CE232965 CRC64;

Query Match 76.7%; Score 33; DB 1; Length 592;

Best Local Similarity 71.4%; Pred. No. 49;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7

DB 34 GRGPAPC 40

RESULT 11

TX32\_PHONI STANDARD; PRT; 32 AA.

ID TX32\_PHONI

AC 076201;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-VAR-2004 (Rel. 43, Last annotation update)

DE Neurotoxin Tx3-2 precursor.

OS Phoneutria nigriventer (Brazilian armed spider).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.

OC NCBI TaxID=6918;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Venom gland;

RX MEDLINE=99053416; PubMed=9839681;

RA Kalapothakis E., Penaforte C.L., Leao R.M., Cruz J.S., Prado V.F., Cordeiro M.N., Diniz C.R., Romano-Silva M.A., Prado M.A.M., Gomez M.V., Beirao P.S.L.;

RT "Cloning, cDNA sequence analysis and patch clamp studies of a toxin from the venom of the armed spider (Phoneutria nigriventer).";

RL Toxicon 36:1971-1980(1998).

RN [2]

RP SEQUENCE OF 38-71.

RC TISSUE=Venom;

RX MEDLINE=93190315; PubMed=8446961;

RA Cordeiro M.N., De Figueiredo S.G., Valentim A.D.C., Diniz C.R., von Eickstedt V.R.D., Gilroy J., Richardson M.;

RT "Purification and amino acid sequences of six Tx3 type neurotoxins from the venom of the Brazilian 'armed' spider Phoneutria nigriventer (Keys).";

RT Toxicon 31:35-42(1993).

RL -!- FUNCTION: Antagonist of L-type calcium channels. Induces immediate clockwise gyration and flaccid paralysis after 6 hours at dose levels of 5 micrograms per mouse.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC -1- SIMILARITY: Belongs to the spider toxin TX3 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/annouance/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL: AFO15664; AAC26168.1; -;  
 DR InterPro: IPR004169; spidertoxin.  
 DR Pfam: PF02819; spidertoxin; 1.  
 KW Calcium channel inhibitor; Toxin; Neurotoxin; Ionic channel inhibitor;  
 KW Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 37  
 FT CHAIN 38 71 NEUROTOXIN TX3-2.  
 FT PROPEP 72 82  
 SQ SEQUENCE 82 AA; 8927 MW; 0CA4E00F5AAAD8B9 CRC64;  
 Query Match 74.4%; Score 32; DB 1; Length 82;  
 Best Local Similarity 71.4%; Pred. No. 8.6;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GRGDSPC 7  
 DB 47 GKGASPC 53  
 RESULT 12  
 KX2\_CAVPO STANDARD; PRT; 239 AA.  
 ID KX2\_CAVPO  
 AC P12323;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glandular kallikrein, prostatic (EC 3.4.21.35) (Tissue kallikrein)  
 DE (Prostate esterase).  
 OS Cavia porcellus (Guinea pig).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 ON NCBI\_TaxID=10141;  
 RN [1]  
 RP MEDLINE=88000549; PubMed=3307909;  
 RA Dunbar J.C., Bradshaw R.A.;  
 RT "Amino acid sequence of guinea pig prostate kallikrein.";  
 RL Biochemistry 26:3471-3478(1987).  
 CC -1- FUNCTION: Glandular kallikreins cleave Met-Lys and Arg-Ser bonds  
 CC in kininogen to release Lys-bradykinin.  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage of Arg-|-Xaa bonds in  
 CC small molecule substrates. Highly selective activator to release  
 CC kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of  
 CC Met-|-Xaa or Leu-|-Xaa.  
 CC -1- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 DR PIR: A27207; A27207.  
 DR HSP; P00757; 1SGF.  
 DR MEROPS: S01.160; -;  
 DR InterPro: IPR009003; Cys Ser trypsin.  
 DR InterPro: IPR001254; Peptidase\_S1.  
 DR InterPro: IPR001314; Peptidase\_S1A.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00020; Tryp\_SPC; 1.  
 DR PROSITE: PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein.  
 FT ACT SITE 41 41 CHARGE RELAY SYSTEM.  
 FT ACT SITE 96 96 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 191 191 CHARGE RELAY SYSTEM.

FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .).  
 FT D-SULFID 7 151 BY SIMILARITY.  
 FT D-SULFID 26 42 BY SIMILARITY.  
 FT D-SULFID 128 197 BY SIMILARITY.  
 FT D-SULFID 162 176 BY SIMILARITY.  
 FT D-SULFID 187 212 BY SIMILARITY.  
 FT VARIANT 50 50 K -> W.  
 SQ SEQUENCE 239 AA; 25989 MW; 56DC81BC10D49A64 CRC64;  
 Query Match 74.4%; Score 32; DB 1; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GDSPC 7  
 DB 208 GDSPC 212  
 RESULT 13  
 FPG\_PSEPK STANDARD; PRT; 270 AA.  
 ID FPG\_PSEPK  
 AC O88C05;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23) (Fapy-DNA  
 DE glycosylase).  
 GN MUTM OR FPG OR PP5125.  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 ON NCBI\_TaxID=160488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L., Seanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
 RA Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,  
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,  
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
 RA Lauber J., Sjepeandic D., Hoheisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Eisen J.A., Timmlis K.N., Duesterhoeft A., Tuemmler B.,  
 RA Fraser C.N.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002)  
 CC -1- FUNCTION: This enzyme may play a significant role in processes  
 CC leading to recovery from mutagenesis and/or cell death by  
 CC alkylating agents (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of DNA containing ring-opened N(7)-  
 CC methylguanine residues, releasing 2,6-diamino-4-hydroxy-5-(N-  
 CC methyl)formamidopyrimidine.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SIMILARITY: Belongs to the FPG family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/annouance/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE016793; AAN70690.1; -;  
 DR TIGR: PP5125; -;  
 DR HAMAP: MF\_00103; -; 1.  
 DR InterPro: IPR000191; Fapy\_DNA\_glyco.  
 DR Pfam: PF01149; Fapy\_DNA\_glyco; 1.  
 DR PROSITE: PS01242; FPG; FALSE NEG.  
 KW DNA repair; Hydrolase; Glycosidase; Zinc; Zinc-finger;  
 KW Complete proteome.

```

FT  ZN FING 245 268 POTENTIAL.
SQ  SEQUENCE 270 AA; 30178 MW; 293D210828C9D26F CRC64;

Query Match 74.4%; Score 32; DB 1; Length 270;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRGDSPC 7
   |||
Db 239 GRGGQPC 245

RESULT 14
FPG_PSESM STANDARD; PRT; 270 AA.
AC Q89AH6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23) (Fapy-DNA
   glycosylase).
GN MUTM OR FPG OR PSPT00414.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=3323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3300;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., Dobay R.T., Durkin A.S., Kolonay J.P.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Serry K., Utterback T.,
RA Vat Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinkour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.V., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
R2 Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -!- FUNCTION: This enzyme may play a significant role in processes
CC leading to recovery from mutagenesis and/or cell death by
CC alkylating agents (By similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of DNA containing ring-opened N(7)-
CC methylguanine residues, releasing 2,6-diamino-4-hydroxy-5-(N-
CC methyl)formamidopyrimidine.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the FPG family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE016857; AAC33958.1; -
CC TIGR; PSPT00414; -
CC HAMAP; MF_00103; -; 1.
CC InterPro; IPR00191; Fapy DNA glyco.
CC Pfam; PF01149; Fapy DNA glyco. 1.
CC PROSITE; PS02442; FPG; FALSE_NEG.
CC DNA repair; Hydrolase; Glycosidase; Zinc; Zinc-finger;
CC Complete proteome.
CC ZN FING 245 268 POTENTIAL.
SQ SEQUENCE 270 AA; 30245 MW; C143DB5BF73054E7 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 270;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GRGDSPC 7
   |||
Db 239 GRGGQPC 245

RESULT 15
MAT2_COCOA STANDARD; PRT; 337 AA.
AC Q9P445;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mating-type protein MAT-2.
GN MAT2.
OS Cochliobolus sativus (Bipolaris sorokiniana).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; Cochliobolus.
OX NCBI_TaxID=45130;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhong S., Steffenson B.J.;
RT "Genetic and molecular characterization of mating type genes in
RT Cochliobolus sativus."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 HMG box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF275374; AAF87724.1; -
CC HSSP; P48436; 1SX9.
CC InterPro; IPR000910; HMG_12_box.
CC Pfam; PF00505; HMG_box; 1.
CC SMART; SM00398; HMG; 1.
CC PROSITE; PS50118; HMG_BOX_2; 1.
CC Transcription regulation; DNA-binding; Nuclear protein; Fertilization.
CC DNA BIND 125 193 HMG_BOX.
SQ SEQUENCE 337 AA; 37582 MW; E83F4A1A8E1EF6CF CRC64;

Query Match 74.4%; Score 32; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDSPC 7
   |||
Db 253 GDSPC 257

Search completed: April 16, 2004, 07:22:16
Job time : 13.25 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: April 16, 2004, 07:03:27 ; Search time 14.7 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-10

Perfect score: 43

Sequence: 1 GRGDSPC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	86.0	704	2 T02902	hypothetical prote
2	35	81.4	1827	2 B70984	probable polyketid
3	34	79.1	228	2 C83612	probable transcrip
4	34	79.1	273	2 A28512	fibronectin - chic
5	34	79.1	285	2 T36296	hypothetical prote
6	34	79.1	441	2 B84264	isochorismate synt
7	34	79.1	525	2 G84406	halolysin [importe
8	34	79.1	691	2 T45933	hypothetical prote
9	34	79.1	1020	2 A23355	fibronectin - chic
10	34	79.1	2265	1 FN90	fibronectin - bovi
11	34	79.1	2386	1 FNH0	fibronectin precu
12	34	79.1	2477	2 S14428	fibronectin precu
13	34	79.1	2481	2 A43908	fibronectin - Afri
14	33	76.7	277	2 E64499	hypothetical prote
15	33	76.7	417	1 E70755	probable phosphos
16	32	74.4	34	2 B43336	neurotoxin rx3-2 -
17	32	74.4	208	2 G96758	protein trypsin in
18	32	74.4	239	2 A27207	tissue kallikrein
19	32	74.4	270	2 T38218	probable abhydrola
20	32	74.4	340	1 S04898	tyb-related protei
21	32	74.4	391	2 AG2709	alanine racemase [
22	32	74.4	391	2 H97491	alanine racemase (
23	32	74.4	427	2 A54714	potassium channel
24	32	74.4	427	2 S43667	inward rectifier p
25	32	74.4	432	2 S71341	inward rectifier p
26	32	74.4	432	2 I52864	potassium channel
27	32	74.4	474	2 JX0301	ubiquinol-cytochro
28	32	74.4	508	2 D81417	probable ATP /GTP
29	32	74.4	546	2 T31255	phtR protein homol

hypothetical prote  
probable pher prot  
parasporal crystal  
parasporal crystal  
parasporal crystal  
parasporal crystal  
parasporal crystal  
parasporal crystal  
parasporal crystal  
hypothetical prote  
carboxypeptidase B  
vinuculin - mouse  
hupK protein - Rhi  
probable aminopept  
RING finger protei  
multidrug resistan

#### ALIGNMENTS

##### RESULT 1

T02902

hypothetical protein T13J8.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999

C:Accession: T02902

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schuel

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z14766

A:Accession: T02902

A:Molecule type: DNA

A:Residues: 1-704 <BEV>

A:Cross-references: EMBL:AL035524

A:Experimental source: cultivar Columbia; BAC clone T13J8

C:Genetics:

A:Map position: 4

A:Note: T13J8.120

Query Match 86.0%; Score 37; DB 2; Length 704;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7

Db 276 RGDSPC 281

##### RESULT 2

B70984

probable polyketide synthetase ppsD [similarity] - Mycobacterium tuberculosis (strain H3

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 03-Nov-2000

C:Accession: B70984

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70984

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1827 <COL>

A:Cross-references: GB:T83857; GB:AL123456; NID:G3242252; PIDN:CAS6093.1; PID:gl781081

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: ppsD

C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac

homology; [acyl-carrier-protein] S-malonyltransferase homology

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F;57-459/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F;566-841/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>  
 F;1439-1618/Domain: short-chain alcohol dehydrogenase homology <SADH>  
 F;1706-1781/Domain: acyl carrier protein homology <ACP>  
 F;1745/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 81.4%; Score 35; DB 2; Length 1827;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 Db 365 GQ3ESPC 372

RESULT 3  
 C83612  
 Probable transcription regulator PA0275 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: C83612  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.-Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A;Reference number: A82950; MUID:20437337; PMID:10934043  
 A;Accession: C93612  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-228 <STO>  
 A;Cross-references: GB:AE004465; GB:AE004091; NID:G5946107; PIDN:AG03664.1; GSPDB:GN001  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA0275

Query Match 79.1%; Score 34; DB 2; Length 228;  
 Best Local Similarity 83.3%; Pred. No. 26;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7  
 |||||  
 Db 46 RGDAPC 51

RESULT 4  
 A28512  
 fibronectin - chicken (fragment)  
 C;Species: Gallus gallus (chicken)  
 C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Aug-1999  
 C;Accession: A28512  
 R;Kubomura, S.; Obara, M.; Karasaki, Y.; Taniguchi, H.; Gotoh, S.; Tsuda, T.; Higashi, K.  
 Biochim. Biophys. Acta 910, 171-181, 1987  
 A;Title: Genetic analysis of the cell binding domain region of the chicken fibronectin gene  
 A;Reference number: A28512; MUID:88050950; PMID:2823899  
 A;Accession: A28512  
 A;Molecule type: DNA  
 A;Residues: 1-273 <KDB>  
 A;Cross-references: GB:X06533; NID:G63393; PIDN:CAA29781.1; PID:G295716  
 A;Note: the authors translated the codon CCG for residue 190 as Gln, CAG for residue 243  
 C;Genetics:  
 A;Intros: 90/1; 129/1; 184/1; 236/1  
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology; fibronectin; fibronectin type III repeat homology; extracellular matrix; glycoprotein; heterodimer; fibronectin type III repeat homology (fragment) <FN3I>  
 F;1-82/Domain: fibronectin type III repeat homology (fragment) <FN3I>  
 F;90-172/Domain: fibronectin type III repeat homology <FN3J>  
 F;167-169/Region: cell attachment (R-G-D) motif  
 F;184-266/Domain: fibronectin type III repeat homology <FN3K>

Query Match 79.1%; Score 34; DB 2; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 6

Db 166 GRGDSP 171  
 |||||

# RESULT 5

T36296  
 Hypothetical protein SCE7.14c - Streptomyces coelicolor  
 C;Species: Streptomyces coelicolor  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 08-Sep-2000  
 C;Accession: T36296  
 R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, May 1998  
 A;Reference number: Z21603  
 A;Accession: T36296  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-285 <SEE>  
 A;Cross-references: EMBL:AL049819; PIDN:CAB42673.1; GSPDB:GN00070; SCOEDB:SCE7.14c  
 A;Experimental source: strain A3(2)  
 C;Genetics:  
 A;Gene: SCOEDB:SCE7.14c  
 C;Superfamily: Streptomyces coelicolor hypothetical protein SCE7.14c

Query Match 79.1%; Score 34; DB 2; Length 285;  
 Best Local Similarity 71.4%; Pred. No. 32;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 Db 215 GRGNQPC 221

# RESULT 6

E84264  
 Isochorismate synthase [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 12-Jun-2003  
 C;Accession: E84264  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable, J.; Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A;Title: Genome sequence of Halobacterium species NRC-1  
 A;Reference number: A84160; MUID:20504483; EMBL:11016950  
 A;Accession: E84264  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-441 <STO>  
 A;Cross-references: GB:AE004437; NID:G10580629; PIDN:AAG19491.1; GSPDB:GN00138  
 C;Genetics:  
 A;Gene: menF  
 C;Superfamily: isochorismate synthase

Query Match 79.1%; Score 34; DB 2; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 6  
 |||||  
 Db 271 GRGDSP 276

# RESULT 7

G84406  
 Halolysin [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
 C;Accession: G84406  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable, J.; Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, E.; Lowe, T.M.; Li  
 A:Title: Genome sequence of *Halobacterium* species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: G84406  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-525 <STO>  
 A:Cross-references: GB:AE004437; NID:gl0581958; PIDN:AG20619.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: hly  
 C:Superfamily: subtilisin; subtilisin homology

Query Match 79.1% Score 34; DB 2; Length 525;  
 Best Local Similarity 85.7%; Pred. No. 60;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
 |||||  
 DB 432 GRGDSNC 438

RESULT 8  
 T45933  
 hypothetical protein F5K20.230 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C:Accession: T45933  
 R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.  
 submitted to the Protein Sequence Database, January 2000  
 A:Reference number: Z33017  
 A:Accession: T45933  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-691 <MON>  
 A:Cross-references: EMBL:AL132960  
 A:Experimental source: cultivar Columbia; BAC clone F5K20  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 87/3; 127/1; 146/3; 175/2; 206/3; 243/1; 267/2; 391/1; 511/3; 594/1; 617/3  
 A:Note: F5K20.230

Query Match 79.1% Score 34; DB 2; Length 691;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 6  
 |||||  
 DB 389 GRGDSPC 394

RESULT 9  
 A29355  
 fibronectin - chicken (fragment)  
 C:Species: *Gallus gallus* (chicken)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 12-Feb-1999  
 C:Accession: A29355  
 R:Norton, P.A.; Hynes, R.O.  
 Mol. Cell. Biol. 7, 4297-4307, 1987  
 A:Title: Alternative splicing of chicken fibronectin in embryos and in normal and trans-  
 A:Reference number: A29355; MUID:88142820; PMID:2830487  
 A:Accession: A29355  
 A:Molecule type: mRNA  
 A:Residues: 1-1020 <NOR>  
 C:Genetics:  
 A:Introns: 176/3  
 C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat  
 C:Keywords: alternative splicing; duplication; extracellular matrix; glycoprotein; heter  
 F1-86/Domain: fibronectin type III repeat homology <FN3>  
 F:92-177/Domain: fibronectin type III repeat homology <FN3>  
 F:180-262/Domain: fibronectin type III repeat homology <FN3>  
 F:257-259/Region: Cell attachment (R-G-D) motif  
 F:274-356/Domain: fibronectin type III repeat homology <FN3K>  
 F:364-446/Domain: fibronectin type III repeat homology <FN3L>

F:454-536/Domain: fibronectin type III repeat homology <FN3M>  
 F:546-628/Domain: fibronectin type III repeat homology <FN3N>  
 F:636-718/Domain: fibronectin type III repeat homology <FN3O>  
 F:837-917/Domain: fibronectin type III repeat homology <FN3P>  
 F:940-979/Domain: fibronectin type I repeat homology <IF10>  
 F:940-969,967-979,985-1012/Disulfide bonds: #status predicted

Query Match 79.1% Score 34; DB 2; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 6  
 |||||  
 DB 256 GRGDSPC 261

RESULT 10  
 FNBO

fibronectin - bovine

C:Species: *Bos primigenius taurus* (cattle)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Oct-2000  
 C:Accession: A26452; B21165; A23292  
 R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.  
 Eur. J. Biochem. 161, 441-453, 1986  
 A:Title: Complete primary structure of bovine plasma fibronectin.  
 A:Reference number: A26452; MUID:87054047; PMID:3780752  
 A:Accession: A26452  
 A:Molecule type: protein  
 A:Residues: 1-2265 <SKO>  
 R:Kornblitt, A.K.; Vibe-Pedersen, K.; Baralle, F.E.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983  
 A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins  
 A:Reference number: A21165; MUID:83221567; PMID:6304699  
 A:Accession: B21165  
 A:Molecule type: mRNA  
 A:Residues: 2170-2265 <KOR>  
 A:Cross-references: GB:K00800; NID:gl63055; PIDN:AAA30521.2; PID:95713323  
 R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottr  
 Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983  
 A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal  
 A:Reference number: A23292; MUID:83117805; PMID:6218503  
 A:Accession: A23292  
 A:Molecule type: protein  
 A:Residues: 1-16, 'C', '18-20', 'S', '22-432-447-463-1367-1517-1567-1673-2062-2176', 'N', '2178-2265  
 C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.  
 C:Comment: The plasma fibronectin molecule consists of two chains, which are connected by  
 C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib  
 aling, and maintenance of cell shape.  
 C:Comment: Plasma fibronectin is synthesized by hepatocytes.  
 C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe  
 C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu  
 F:21-241/Domain: fibrin and heparin binding <FBR>  
 F:21-56/Domain: fibronectin type I repeat homology <IF1>  
 F:66-104/Domain: fibronectin type I repeat homology <IF2>  
 F:110-148/Domain: fibronectin type I repeat homology <IF3>  
 F:155-194/Domain: fibronectin type I repeat homology <IF4>  
 F:200-239/Domain: fibronectin type I repeat homology <IF5>  
 F:277-577/Domain: collagen binding <CBR>  
 F:277-311/Domain: fibronectin type I repeat homology <IF6>  
 F:329-370/Domain: fibronectin type II repeat homology <2F1>  
 F:389-430/Domain: fibronectin type II repeat homology <2F2>  
 F:439-477/Domain: fibronectin type I repeat homology <IF7>  
 F:487-524/Domain: fibronectin type I repeat homology <IF8>  
 F:530-568/Domain: fibronectin type I repeat homology <IF9>  
 F:578-661/Domain: fibronectin type III repeat homology <FN3A>  
 F:688-770/Domain: fibronectin type III repeat homology <FN3B>  
 F:779-860/Domain: fibronectin type III repeat homology <FN3C>  
 F:875-957/Domain: fibronectin type III repeat homology <FN3D>  
 F:965-1046/Domain: fibronectin type III repeat homology <FN3E>  
 F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>  
 F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>  
 F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>  
 F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>

F:1410-1517/Domain: cell attachment <CAD>  
 F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>  
 F:1493-1495/Region: cell attachment (R-G-D) motif  
 F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>  
 F:1600-1870/Domain: heparin binding <HB2>  
 F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>  
 F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>  
 F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>  
 F:1970-1972/Region: cell attachment (R-G-D) motif  
 F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>  
 F:1985-2216/Domain: fibrin binding <FB2>  
 F:2085-2124/Domain: fibronectin type I repeat homology <IF1J>  
 F:2130-2167/Domain: fibronectin type I repeat homology <IF1L>  
 F:2174-2209/Domain: fibronectin type I repeat homology <IF1I>  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:3/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of fibrin) #status experimental  
 F:121-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-223,227-239,277-304,302-3  
 F:2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted  
 F:399,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent  
 F:1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental  
 F:2246/Disulfide bonds: interchain (to 2250) #status predicted  
 F:2250/Disulfide bonds: interchain (to 2246) #status predicted  
 F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 79.1%; Score 34; DB 1; Length 2265;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 3;  
 QY 1 GRGDSF 6  
 DB 1492 GRGDSF 1497

RESULT 11  
 FNHU  
 fibronectin precursor [validated] - human  
 N/A:ternate names: fibronectin splice form ED-A  
 C/Species: Homo sapiens (man)  
 C/Date: 27-Nov-1995 #sequence revision 31-Mar-1993 #text change C8-Dec-2000  
 C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22  
 R/Dean, D.C.; Bowles, C.L.; Bourgeois, S.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987  
 A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.  
 A/Reference number: A26460; MUID:87175578; PMID:3331656  
 A/Accession: A26460  
 A/Molecule type: DNA  
 A/Residues: 1-49 <DEA>  
 A/Cross-references: GB:M15801; NID:g182686; PIDN:AAA53376.1; PID:g553293  
 R/Oldberg, A.; Ruoslahti, E.  
 J. Biol. Chem. 261, 2113-2116, 1986  
 A/Title: Evolution of the fibronectin gene.  
 A/Reference number: A26284; MUID:86111901; PMID:3003095  
 A/Accession: A26284  
 A/Molecule type: DNA  
 A/Residues: 1447-1540 <OLD>  
 A/Cross-references: GB:M12549; NID:g182688  
 A/Note: the authors translated the codon TTC for residue 1434 as Glu  
 R/Paolella, G.; Henchcliffe, C.; Sebastiao, G.; Baralle, F.E.  
 Nucleic Acids Res. 16, 3545-3557, 1988  
 A/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B  
 A/Reference number: S00848; MUID:88233940; PMID:3375063  
 A/Accession: S03917  
 A/Molecule type: DNA  
 A/Residues: 1594-1767, 'V', 1769-1793 <PAO>  
 A/Cross-references: EMBL:X07118; NID:g31402  
 A/Note: the authors translated the codon AAC for residue 1631 as Asp  
 R/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.  
 FEBS Lett. 207, 287-291, 1986  
 A/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.  
 A/Reference number: A24854; MUID:87030929; PMID:3770201  
 A/Accession: A24854  
 A/Molecule type: DNA

A/Residues: 1992-2147 <VIB>  
 A/Cross-references: GB:X04530; NID:g31436  
 R/Gutman, A.; Yanada, K.M.; Kornblitt, A.  
 FEBS Lett. 207, 145-148, 1986  
 A/Title: Human fibronectin is synthesized as a pre-propolypeptide.  
 A/Reference number: A24476; MUID:87030890; PMID:3770189  
 A/Accession: A24476  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-14, 'Q', 16-38 <GUT>  
 R/Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.  
 EMBO J. 4, 1755-1759, 1985  
 A/Title: Primary structure of human fibronectin: differential splicing may generate at  
 A/Reference number: A91008; MUID:85284965; PMID:2992939  
 A/Accession: A91008  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 32-1344, 1346-2080; 2112-2386 <KOR>  
 A/Cross-references: GB:X02761  
 R/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
 Nucleic Acids Res. 12, 5853-5868, 1984  
 A/Title: Human fibronectin: cell specific alternative mRNA splicing generates polypepti  
 A/Reference number: A93529; MUID:84272258; PMID:6462519  
 A/Accession: A93529  
 A/Molecule type: mRNA  
 A/Residues: 973-2080; 2112-2386 <KO2>  
 A/Cross-references: GB:X00739  
 R/Oldberg, A.; Linney, E.; Ruoslahti, E.  
 J. Biol. Chem. 258, 10193-10196, 1983  
 A/Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell  
 A/Reference number: A21011; MUID:83290929; PMID:6688418  
 A/Accession: A21011  
 A/Molecule type: mRNA  
 A/Residues: 1434-1537 <OL2>  
 A/Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683  
 R/Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.  
 Biochemistry 24, 2698-2704, 1985  
 A/Title: Human cellular fibronectin: comparison of the carboxy-terminal portion with r  
 A/Reference number: A90495; MUID:85280409; PMID:2992573  
 A/Accession: A90495  
 A/Molecule type: mRNA  
 A/Residues: 1594-2386 <BER>  
 A/Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697  
 R/Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.  
 FEBS Lett. 186, 31-34, 1985  
 A/Title: Isolation and characterization of cDNA clones for human liver fibronectin.  
 A/Reference number: A22245; MUID:85231203; PMID:2989004  
 A/Accession: A22245  
 A/Molecule type: mRNA  
 A/Residues: 1948-2067 <JME>  
 A/Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706  
 A/Accession: B22245  
 A/Molecule type: mRNA  
 A/Residues: 1975-1991; 2017-2039 <JME>  
 A/Cross-references: GB:M27590  
 R/Sekiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.  
 Biochemistry 25, 4936-4941, 1986  
 A/Title: Human liver fibronectin complementary DNAs: identification of two different me  
 A/Reference number: I52394; MUID:87026578; PMID:3021206  
 A/Accession: I52394  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1978-1990, 2016-2018, 'N', 2020-2081; 2113-2127 <SEK>  
 A/Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704  
 R/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983  
 A/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin  
 A/Reference number: A21165; MUID:83221567; PMID:6304699  
 A/Accession: A21165  
 A/Molecule type: mRNA  
 A/Residues: 2291-2386 <KO3>  
 A/Cross-references: GB:X00793; NID:g182681; PIDN:AAA52463.1; PID:g182684  
 R/Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.

J. Biol. Chem. 258, 12670-12674, 1983  
 A;Title: Primary structure of human plasma fibronectin.  
 A;Reference number: A92398; XUID:84032463; PMID:6630202  
 A;Accession: A92398  
 A;Molecule type: protein  
 A;Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>  
 R;Garcia-Pardo, A.; Gold, L.I.  
 Arch. Biochem. Biophys. 304, 181-188, 1993  
 A;Title: Further characterization of the binding of fibronectin to gelatin reveals the  
 A;Reference number: S34791; XUID:93312301; PMID:8323285  
 A;Accession: S34791  
 A;Molecule type: protein  
 A;Residues: 291-300;551-560 <GAR2>  
 R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.  
 Thromb. Res. 43, 469-477, 1986  
 A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.  
 A;Reference number: A60904; XUID:87019725; PMID:3532418  
 A;Accession: A60904  
 A;Molecule type: protein  
 A;Residues: 293-301 <GRI>  
 R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.  
 J. Biol. Chem. 260, 12136-12141, 1985  
 A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl  
 A;Reference number: A23901; XUID:86038277; PMID:3900070  
 A;Accession: A23901  
 A;Molecule type: protein  
 A;Residues: 616-677, 'C', 679-703, 'PT' <CAL>  
 R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.  
 J. Biol. Chem. 257, 9593-9597, 1982  
 A;Title: The cell attachment domain of fibronectin. Determination of the primary structu  
 A;Reference number: A92386; XUID:82265604; PMID:7050098  
 A;Accession: A92386  
 A;Molecule type: protein  
 A;Residues: 1441-1548 <PIE>  
 A;Note: residues 1524-1527 are responsible for the cell-binding activity  
 R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.  
 Biochem. J. 241, 923-928, 1987  
 A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom  
 A;Reference number: A32517; XUID:87241275; PMID:3593230  
 A;Accession: A32517  
 A;Molecule type: protein  
 A;Residues: 1589-1630, 'T', 1722-2081, 2113-2244 <TRE>  
 R;Fressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand  
 Biochem. J. 274, 731-738, 1991  
 A;Title: Human plasma fibronectin. Demonstration of structural differences between the A  
 A;Reference number: S14357; XUID:9150085; PMID:2012601  
 A;Accession: S14357  
 A;Molecule type: protein  
 A;Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>  
 R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.  
 J. Biol. Chem. 260, 10320-10325, 1985  
 A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal  
 A;Reference number: A23891; XUID:85261459; PMID:4019516  
 A;Accession: A23891  
 A;Molecule type: protein  
 A;Residues: 2071-2080;2112-2356 <GAR4>  
 C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis  
 ation, and transformation.  
 C;Genetics:  
 A;Gene: GDB: FNI  
 A;Cross-references: GDB:119135; OMIM:135600  
 A;Map position: 2Q34-2Q34  
 A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1  
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe  
 C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat  
 F;1-26/Domain: signal sequence #status predicted <SIG>  
 F;27-31/Domain: propeptide #status predicted <PRO>  
 F;32-2386/Product: fibronectin #status experimental <MAT>  
 F;52-272/Domain: fibrin and heparin binding <PHB>  
 F;52-87/Domain: fibronectin type I repeat homology <1F1>  
 F;97-135/Domain: fibronectin type I repeat homology <1F2>  
 F;141-179/Domain: fibronectin type I repeat homology <1F3>

F;186-225/Domain: fibronectin type I repeat homology <1F4>  
 F;231-270/Domain: fibronectin type I repeat homology <1F5>  
 F;308-608/Domain: collagen binding <CB>  
 F;308-342/Domain: fibronectin type I repeat homology <1F6>  
 F;360-461/Domain: fibronectin type II repeat homology <2F1>  
 F;420-461/Domain: fibronectin type II repeat homology <2F2>  
 F;470-588/Domain: fibronectin type I repeat homology <1F7>  
 F;518-559/Domain: fibronectin type I repeat homology <1F8>  
 F;561-599/Domain: fibronectin type I repeat homology <1F9>  
 F;609-692/Domain: fibronectin type III repeat homology <3FA>  
 F;616-706/Domain: heparin binding <HPB>  
 F;719-801/Domain: fibronectin type III repeat homology <3FB>  
 F;810-891/Domain: fibronectin type III repeat homology <3FC>  
 F;906-988/Domain: fibronectin type III repeat homology <3FD>  
 F;996-1077/Domain: fibronectin type III repeat homology <3FE>  
 F;1086-1164/Domain: fibronectin type III repeat homology <3FF>  
 F;1173-1258/Domain: fibronectin type III repeat homology <3FG>  
 F;1266-1349/Domain: fibronectin type III repeat homology <3FH>

Query Match 79.1%; Score 34; DB 1; Length 2386;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 3;

QY 1 GRGDSP 6  
 Db 1523 GRGDSP 1528

## RESULT 12

S14428  
 fibronectin precursor - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 20-Aug-1999  
 C;Accession: S14428; S12455; A22319; S46203; S00459; A2752; I59049  
 R;Hynes, R.O.  
 submitted to the EMBL Data Library, July 1989  
 A;Reference number: S14428  
 A;Accession: S14428  
 A;Molecule type: mRNA  
 A;Residues: 1-2477 <HYN>  
 A;Cross-references: EMBL:X15906; NID:956163; PIDN:CA34020.1; PID:956164  
 R;Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.  
 EMBO J. 6, 2573-2580, 1987  
 A;Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.  
 A;Reference number: S12455; XUID:88054951; PMID:2445560  
 A;Accession: S12455  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 609-1810, 'T', 1812-2283 <SCH>  
 A;Cross-references: EMBL:X15906  
 R;Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984  
 A;Title: A single rat fibronectin gene generates three different mRNAs by alternative sp  
 A;Reference number: A22319; XUID:84298097; PMID:6089177  
 A;Accession: A22319  
 A;Molecule type: DNA  
 A;Residues: 2052-2237 <TAM>  
 R;Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstrøm, B.  
 Biochem. J. 301, 745-751, 1994  
 A;Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in  
 A;Reference number: S46203; XUID:94330948; PMID:7519849  
 A;Accession: S46203  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1183-1192, 'GN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PY', 1385-1399 <FAL>  
 R;Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.  
 EMBO J. 6, 2565-2572, 1987  
 A;Title: Organization of the fibronectin gene provides evidence for exon shuffling durin  
 A;Reference number: S00459; XUID:88054950; PMID:3113323  
 A;Accession: S00459  
 A;Molecule type: DNA  
 A;Residues: 1-139, 2382-2477 <PAT>  
 A;Cross-references: EMBL:X05831



A>Note: the authors translated the codon CCG for residues 51 and 94 as Ala  
 R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.  
 Cell 35, 421-431, 1983  
 A>Title: Three different fibronectin mRNAs arise by alternative splicing within the coding region  
 A:Reference number: A27252; MUID:84082067; PMID:6317187  
 A:Accession: A27252  
 A:Molecule type: mRNA  
 A:Residues: 1596-1720, T', 1722, 1813-2477 <SC>  
 R:Odermatt, E.; Tamkun, J.W.; Hynes, R.O.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985  
 A>Title: Repeating modular structure of the fibronectin gene: Relationship to protein structure  
 A:Reference number: I59049; MUID:86016741; PMID:3863113  
 A:Accession: I59049  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1722-1810 <RRS>  
 A:Cross-references: GB:M11750; NID:9284164; PIDN:AAA41170.1; PID:9554437  
 C:Genetics:  
 A:Introns: 51/1; 94/1; 2416/3; 2454/3  
 C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology  
 C:Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; duplication  
 F:1-32/Domain: signal sequence #status predicted <SIG>  
 F:33-2477/Product: fibronectin #status predicted <MAT>  
 F:53-88/Domain: fibronectin type I repeat homology <1F1>  
 F:98-136/Domain: fibronectin type I repeat homology <1F2>  
 F:142-180/Domain: fibronectin type I repeat homology <1F3>  
 F:187-226/Domain: fibronectin type I repeat homology <1F4>  
 F:232-271/Domain: fibronectin type I repeat homology <1F5>  
 F:308-342/Domain: fibronectin type I repeat homology <1F6>  
 F:360-401/Domain: fibronectin type II repeat homology <2F1>  
 F:420-461/Domain: fibronectin type II repeat homology <2F2>  
 F:470-508/Domain: fibronectin type I repeat homology <1F7>  
 F:518-555/Domain: fibronectin type I repeat homology <1F8>  
 F:561-599/Domain: fibronectin type I repeat homology <1F9>  
 F:609-692/Domain: fibronectin type III repeat homology <FN3A>  
 F:718-800/Domain: fibronectin type III repeat homology <FN3B>  
 F:809-890/Domain: fibronectin type III repeat homology <FN3C>  
 F:905-987/Domain: fibronectin type III repeat homology <FN3D>  
 F:995-1076/Domain: fibronectin type III repeat homology <FN3E>  
 F:1085-1164/Domain: fibronectin type III repeat homology <FN3F>  
 F:1172-1257/Domain: fibronectin type III repeat homology <FN3G>  
 F:1265-1348/Domain: fibronectin type III repeat homology <FN3H>  
 F:1356-1439/Domain: fibronectin type III repeat homology <FN3I>  
 F:1447-1529/Domain: fibronectin type III repeat homology <FN3J>  
 F:1537-1619/Domain: fibronectin type III repeat homology <FN3K>  
 F:1614-1616/Region: cell attachment (R-G-D) motif  
 F:1631-1713/Domain: fibronectin type III repeat homology <FN3L>  
 F:1721-1803/Domain: fibronectin type III repeat homology <FN3M>  
 F:1811-1893/Domain: fibronectin type III repeat homology <FN3N>  
 F:1903-1984/Domain: fibronectin type III repeat homology <FN3O>  
 F:1992-2074/Domain: fibronectin type III repeat homology <FN3P>  
 F:2181-2183/Region: cell attachment (R-G-D) motif  
 F:2193-2273/Domain: fibronectin type III repeat homology <FN3Q>  
 F:2296-2335/Domain: fibronectin type I repeat homology <1F10>  
 F:2341-2378/Domain: fibronectin type I repeat homology <1F11>  
 F:2385-2420/Domain: fibronectin type I repeat homology <1F12>  
 F:53-79, 77-88, 98-126, 124-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333-368, 2366-2378, 2385-2411, 2409-2420/Disulfide bonds: #status predicted  
 F:2458/Disulfide bonds: interchain (to 2462) #status predicted  
 F:2462/Disulfide bonds: interchain (to 2458) #status predicted

Query Match 79.1%; Score 34; DB 2; Length 2477;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6

Db -613 GRGDSP 16:8

RESULT 13

A43908

fibronectin - African clawed frog

C:Species: Xenopus laevis (African clawed frog)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 12-Feb-1999  
 C:Accession: A43908  
 R:Desimone, D.W.; Norton, P.A.; Hynes, R.O.  
 Dev. Biol. 149, 357-369, 1992  
 A>Title: Identification and characterization of alternatively spliced fibronectin mRNAs  
 A:Reference number: A43908; MUID:92111942; PMID:1730390  
 A:Accession: A43908  
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-2481 <DES>  
 A:Cross-references: GB:M77820

A:Note: sequence extracted from NCBI backbone (NCBI:77473)  
 C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology  
 C:Keywords: duplication; extracellular matrix; glycoprotein; heterodimer  
 F:55-90/Domain: fibronectin type I repeat homology <1F1>  
 F:100-138/Domain: fibronectin type I repeat homology <1F2>  
 F:144-182/Domain: fibronectin type I repeat homology <1F3>  
 F:189-228/Domain: fibronectin type I repeat homology <1F4>  
 F:234-273/Domain: fibronectin type I repeat homology <1F5>  
 F:309-343/Domain: fibronectin type I repeat homology <1F6>  
 F:361-402/Domain: fibronectin type II repeat homology <2F1>  
 F:421-462/Domain: fibronectin type II repeat homology <2F2>  
 F:471-509/Domain: fibronectin type I repeat homology <1F7>  
 F:519-556/Domain: fibronectin type I repeat homology <1F8>  
 F:562-600/Domain: fibronectin type I repeat homology <1F9>  
 F:610-693/Domain: fibronectin type III repeat homology <FN3A>  
 F:719-801/Domain: fibronectin type III repeat homology <FN3B>  
 F:810-891/Domain: fibronectin type III repeat homology <FN3C>  
 F:906-988/Domain: fibronectin type III repeat homology <FN3D>  
 F:996-1077/Domain: fibronectin type III repeat homology <FN3E>  
 F:1086-1165/Domain: fibronectin type III repeat homology <FN3F>  
 F:1173-1258/Domain: fibronectin type III repeat homology <FN3G>  
 F:1266-1349/Domain: fibronectin type III repeat homology <FN3H>  
 F:1357-1440/Domain: fibronectin type III repeat homology <FN3I>  
 F:1448-1530/Domain: fibronectin type III repeat homology <FN3J>  
 F:1538-1620/Domain: fibronectin type III repeat homology <FN3K>  
 F:1615-1617/Region: cell attachment (R-G-D) motif  
 F:1632-1714/Domain: fibronectin type III repeat homology <FN3L>  
 F:1722-1804/Domain: fibronectin type III repeat homology <FN3M>  
 F:1812-1894/Domain: fibronectin type III repeat homology <FN3N>  
 F:1904-1985/Domain: fibronectin type III repeat homology <FN3O>  
 F:1993-2075/Domain: fibronectin type III repeat homology <FN3P>  
 F:2197-2277/Domain: fibronectin type III repeat homology <FN3Q>  
 F:2301-2340/Domain: fibronectin type I repeat homology <1F10>  
 F:2346-2383/Domain: fibronectin type I repeat homology <1F11>  
 F:2390-2425/Domain: fibronectin type I repeat homology <1F12>  
 F:55-81, 79-90, 100-128, 126-138, 144-172, 170-182, 189-218, 216-228, 234-263, 261-273, 309-336, 333-368, 2373, 2371-2383, 2390-2416, 2414-2425/Disulfide bonds: #status predicted  
 F:2459/Disulfide bonds: interchain (to 2463) #status predicted  
 F:2463/Disulfide bonds: interchain (to 2459) #status predicted

Query Match 79.1%; Score 34; DB 2; Length 2481;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSP 6

Db 1614 GRGDSP 1619

RESULT 14

F64499

hypothetical protein MJ1599 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: F64499

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: F64499  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-277 <BUL>  
A:Cross-references: GB:U67600; GB:L77117; NID:G2826437; PIDN:AB99626.1; PID:G1500496; T  
C:Genetics:  
A:Map position: REV1572925-1572092

Query Match 76.7%; Score 33; DB 2; Length 277;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGDSPC 7  
DB 131 RGDSPC 136

## RESULT 15

E70755  
probable phosphoesterase (EC 3.1.1.-) Rv1277 [similarity] - Mycobacterium tuberculosis  
C:Species: Mycobacterium tuberculosis  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: E70755  
R: Cole, S.C.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
; Connor, K.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70755  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-417 <COL>  
A:Cross-references: GB:Z77137; GB:AL123456; NID:G3261593; PIDN:CA500917.1; PID:e1299886;  
A:Experimental source: strain H37RV  
C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.  
C:Genetics:  
A:Gene: Rv1277  
C:Superfamily: Mycobacterium tuberculosis probable phosphoesterase Rv1277; phosphoesterase  
C:Keywords: hydrolase  
F:20-122/Domain: phosphoesterase core homology <PEC>

Query Match 76.7%; Score 33; DB 1; Length 417;  
Best Local Similarity 71.4%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
DB 9 GRGPAPC 15

Search completed: April 16, 2004, 07:23:25  
Job time : 15.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:28:08 ; Search time: 47.075 Seconds  
(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588b-10

Perfect score: 43

Sequence: 1 GRGSPC 7

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	10	US-09-911-569-8
2	43	100.0	7	11	US-09-991-588b-10
3	43	100.0	7	14	US-10-200-879-8
4	43	100.0	9	9	US-09-892-071-11
5	43	100.0	10	9	US-09-893-292-1
6	43	100.0	10	9	US-09-892-071-1
7	43	100.0	10	9	US-09-892-071-8
8	43	100.0	10	13	US-10-007-270-37
9	43	100.0	22	14	US-09-911-569-120
10	43	100.0	22	14	US-10-200-879-120
11	43	100.0	25	10	US-09-911-569-118
12	43	100.0	25	10	US-09-911-569-119
13	43	100.0	25	14	US-10-200-879-118
14	43	100.0	25	14	US-10-200-879-119
15	43	100.0	96	10	US-09-911-569-62

16	43	100.0	96	14	US-10-200-879-62	Sequence 62, Appl
17	40	93.0	477	12	US-10-425-114-71233	Sequence 71233, A
18	39	90.7	83	12	US-10-424-599-191403	Sequence 191403, A
19	37	86.0	6	10	US-09-911-569-14	Sequence 14, Appl
20	37	86.0	6	14	US-10-200-879-14	Sequence 14, Appl
21	37	86.0	16	10	US-09-911-569-105	Sequence 105, App
22	37	86.0	16	14	US-10-200-879-105	Sequence 105, App
23	37	86.0	30	10	US-09-911-569-100	Sequence 100, App
24	37	86.0	30	14	US-10-200-879-100	Sequence 100, App
25	37	86.0	95	10	US-09-911-569-63	Sequence 63, Appl
26	37	86.0	95	14	US-10-200-879-63	Sequence 63, Appl
27	35	81.4	85	12	US-10-276-774-1672	Sequence 1672, Ap
28	35	81.4	87	12	US-10-424-599-147277	Sequence 147277, A
29	35	81.4	96	12	US-10-425-114-61346	Sequence 61346, A
30	35	81.4	238	12	US-10-425-114-61346	Sequence 61346, A
31	35	81.4	339	12	US-10-425-114-64888	Sequence 64888, A
32	35	81.4	1827	9	US-09-712-363-261	Sequence 261, App
33	34	79.1	6	9	US-09-837-969A-28	Sequence 28, Appl
34	34	79.1	6	9	US-09-837-969A-46	Sequence 46, Appl
35	34	79.1	6	9	US-09-888-260-15	Sequence 15, Appl
36	34	79.1	6	9	US-09-961-834-1	Sequence 1, Appli
37	34	79.1	6	9	US-09-364-597A-21	Sequence 21, Appl
38	34	79.1	6	9	US-09-177-843-1	Sequence 1, Appli
39	34	79.1	6	9	US-09-841-321A-28	Sequence 28, Appl
40	34	79.1	6	9	US-09-841-321A-46	Sequence 46, Appl
41	34	79.1	6	9	US-09-871-974-3	Sequence 3, Appli
42	34	79.1	6	9	US-09-972-772-31	Sequence 31, Appl
43	34	79.1	6	11	US-09-991-588B-7	Sequence 7, Appli
44	34	79.1	6	12	US-09-961-381A-3	Sequence 3, Appli
45	34	79.1	6	12	US-09-380-682-63	Sequence 63, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-911-569-8  
; Sequence 8, Application US/09911569  
; Publication NO. US20030069173A1

#### GENERAL INFORMATION:

APPLICANT: LANLEY-NELSON, PAMELA

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/911,569

FILING DATE: 23-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
 US-09-911-569-8

Query Match 100.0%; Score 43; DB 10; Length 7;  
 Best Local Similarity 100.0%; Pred. No. le+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7  
 Db 1 GRGDSPC 7

RESULT 2  
 US-09-991-588B-10  
 ; Sequence 10, Application US/09991588B  
 ; Publication No. US20030219429A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Budny, John A.  
 ; TITLE OF INVENTION: Composition and Method for Bone Regeneration  
 ; FILE REFERENCE: 1008-120, US  
 ; CURRENT APPLICATION NUMBER: US/09/991,588B  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 09/122,346  
 ; PRIOR FILING DATE: 1998-07-24  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SEQ ID NO 10  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Purchased commercially or sequence is synthesized  
 US-09-991-588B-10

Query Match 100.0%; Score 43; DB 11; Length 7;  
 Best Local Similarity 100.0%; Pred. No. le+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7  
 Db 1 GRGDSPC 7

RESULT 3  
 US-10-200-879-8  
 ; Sequence 8, Application US/10200879  
 ; Publication No. US2003014230A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAWLEY-NELSON, PAMELA  
 ; INVENTOR: LAM, JIANQING  
 ; SHIH, POJEN  
 ; JESSE, JOEL A.  
 ; SCHIFFERLI, KEVIN P.  
 ; GEBEYEHU, GULILAT  
 ; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
 ; NUMBER OF SEQUENCES: 120  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
 ; STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
 ; CITY: BOULDER  
 ; STATE: CO  
 ; COUNTRY: US  
 ; ZIP: 80303  
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/200,879  
 FILING DATE: 23-Jul-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 09/911,569  
 FILING DATE: 23-JUL-2001  
 APPLICATION NUMBER: US 09/039,780  
 FILING DATE: 16-MAR-1998  
 APPLICATION NUMBER: US 08/818,200  
 FILING DATE: 14-MAR-1997  
 APPLICATION NUMBER: US 08/658,130  
 FILING DATE: 04-JUN-1996  
 APPLICATION NUMBER: US 08/477,354  
 FILING DATE: 07-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: SULLIVAN, SALLY A.  
 REGISTRATION NUMBER: 32,064  
 REFERENCE/DOCKET NUMBER: 32-95E  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303)499-8080  
 TELEFAX: (303)499-8089  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
 US-10-200-879-8

Query Match 100.0%; Score 43; DB 14; Length 7;  
 Best Local Similarity 100.0%; Pred. No. le+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7  
 Db 1 GRGDSPC 7

RESULT 4  
 US-09-892-071-11  
 ; Sequence 11, Application US/09892071  
 ; Patent No. US20020028913A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pierschbacher, Michael D.  
 ; APPLICANT: Ruoslahti, Erkki I.  
 ; TITLE OF INVENTION: Conformationally Stabilized Cell  
 ; TITLE OF INVENTION: Adhesion Peptides  
 ; FILE REFERENCE: P-LA 3637  
 ; CURRENT APPLICATION NUMBER: US/09/892,071  
 ; CURRENT FILING DATE: 2001-06-26  
 ; PRIOR APPLICATION NUMBER: US 09/366,991  
 ; PRIOR FILING DATE: 1999-08-04  
 ; PRIOR APPLICATION NUMBER: US 08/124,992  
 ; PRIOR FILING DATE: 1993-09-21  
 ; PRIOR APPLICATION NUMBER: US 08/048,576  
 ; PRIOR FILING DATE: 1993-04-15  
 ; PRIOR APPLICATION NUMBER: US 07/803,797  
 ; PRIOR FILING DATE: 1991-12-06  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: 2
; OTHER INFORMATION: Xaa-Penicillamine (Pen)
US-09-892-071-11
Query Match      100.0%; Score 43; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 5
US-09-809-292-1
; Sequence 1, Application US/03809292
; Patent No. US20020022264A1
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Sean M.
; APPLICANT: Hofland, Hans
; TITLE OF INVENTION: Delivery Vehicles Comprising Stable Lipid/Nucleic Acid
; TITLE OF INVENTION: Complexes
; FILE REFERENCES: Targeting Ligand
; CURRENT APPLICATION NUMBER: US/09/809,292
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 08/652,018
; PRIOR FILING DATE: 1996-05-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide ligand
; OTHER INFORMATION: for integrin receptor
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(10)
; OTHER INFORMATION: Cyclic Peptide for targeting integrin receptor.
; NAME/KEY: PEPTIDE
; LOCATION: (2)
; OTHER INFORMATION: Xaa: penicillamine
US-09-809-292-1
Query Match      100.0%; Score 43; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 6
US-09-892-071-1
; Sequence 1, Application US/09892071
; Patent No. US20020028913A1
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Ruoslahti, Erkki I.
; TITLE OF INVENTION: Conformationally Stabilized Cell
; TITLE OF INVENTION: Adhesion Peptides
; FILE REFERENCE: P-LA 3637
; CURRENT APPLICATION NUMBER: US/09/892,071
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 09/365,991
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 08/124,992
; PRIOR FILING DATE: 1993-09-21
; PRIOR APPLICATION NUMBER: US 08/048,576
; PRIOR FILING DATE: 1993-04-15
; PRIOR APPLICATION NUMBER: US 07/803,797
; PRIOR FILING DATE: 1991-12-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: 2
; OTHER INFORMATION: Xaa-Penicillamine (Pen)
US-09-892-071-8
Query Match      100.0%; Score 43; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 8
US-10-007-270-37
; Sequence 37, Application US/10007270
; Publication No. US20020160954A1

```

```

; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: 2
; OTHER INFORMATION: Xaa-Penicillamine (Pen)
US-09-892-071-11
Query Match      100.0%; Score 43; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 5
US-09-809-292-1
; Sequence 1, Application US/03809292
; Patent No. US20020022264A1
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Sean M.
; APPLICANT: Hofland, Hans
; TITLE OF INVENTION: Delivery Vehicles Comprising Stable Lipid/Nucleic Acid
; TITLE OF INVENTION: Complexes
; FILE REFERENCES: Targeting Ligand
; CURRENT APPLICATION NUMBER: US/09/809,292
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 08/652,018
; PRIOR FILING DATE: 1996-05-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide ligand
; OTHER INFORMATION: for integrin receptor
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(10)
; OTHER INFORMATION: Cyclic Peptide for targeting integrin receptor.
; NAME/KEY: PEPTIDE
; LOCATION: (2)
; OTHER INFORMATION: Xaa: penicillamine
US-09-809-292-1
Query Match      100.0%; Score 43; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 6
US-09-892-071-1
; Sequence 1, Application US/09892071
; Patent No. US20020028913A1
; GENERAL INFORMATION:
; APPLICANT: Pierschbacher, Michael D.
; APPLICANT: Ruoslahti, Erkki I.
; TITLE OF INVENTION: Conformationally Stabilized Cell
; TITLE OF INVENTION: Adhesion Peptides
; FILE REFERENCE: P-LA 3637
; CURRENT APPLICATION NUMBER: US/09/892,071
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 09/365,991
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 08/124,992
; PRIOR FILING DATE: 1993-09-21
; PRIOR APPLICATION NUMBER: US 08/048,576
; PRIOR FILING DATE: 1993-04-15
; PRIOR APPLICATION NUMBER: US 07/803,797
; PRIOR FILING DATE: 1991-12-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: 2
; OTHER INFORMATION: Xaa-Penicillamine (Pen)
US-09-892-071-8
Query Match      100.0%; Score 43; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRGDSPC 7
    |||||
Db 3 GRGDSPC 9

RESULT 8
US-10-007-270-37
; Sequence 37, Application US/10007270
; Publication No. US20020160954A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Hageman, Gregory S.
; APPLICANT: Kuehn, Markus H.
; APPLICANT: University of Iowa Research Foundation
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR OCULAR ABNORMALITIES
; FILE REFERENCE: 020618-020120US
; CURRENT APPLICATION NUMBER: US/10/007,270
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 09/430,195
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/183,972
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Inhibitor
; NAME/KEY: MOD RES
; LOCATION: (2)
; OTHER INFORMATION: Xaa is penicillamine
US-10-007-270-37

Query Match      100.0%; Score 43; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRGDSPC 7
DB      3 GRGDSPC 9

RESULT 9
US-09-911-569-120
; Sequence 120, Application US/09911569
; Publication No. US20030069173A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NELSON, PAMELA
; SHIH, POJEN
; LAN, JIANQING
; JESSE, JOEL A.
; SCHIFFERLI, KEVIN P.
; GEBYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/911,569
; FILING DATE: 23-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080

US-09-911-569-120
; Sequence 120, Application US/10200879
; Publication No. US20030144230A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NELSON, PAMELA
; LAN, JIANQING
; SHIH, POJEN
; JESSE, JOEL A.
; SCHIFFERLI, KEVIN P.
; GEBYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/200,879
; FILING DATE: 23-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/911,569
; FILING DATE: 23-JUL-2001
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; APPLICATION NUMBER: US 08/818,200
; FILING DATE: 14-MAR-1997
; APPLICATION NUMBER: US 08/658,130
; FILING DATE: 04-JUN-1996
; APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid

```

STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 123;  
US-10-200-879-120

Query Match 100.0%; Score 43; DB 14; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.85; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
DB 15 GRGDSPC 21

RESULT 11  
US-09-911-569-118  
; Sequence 118, Application US/09911569  
; Publication No. US20030069173A1  
; GENERAL INFORMATION:  
; APPLICANT: HAWLEY-NELSON, PAMELA  
; LAN, JIANQING  
; SHIH, POJEN  
; JESSE, JOEL A.  
; SCHIFFERLI, KEVIN P.  
; GEBYERHU, GULLIAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/911,569  
FILING DATE: 23-Jul-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95D  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 118:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 118;  
US-09-911-569-118

Query Match 100.0%; Score 43; DB 10; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.95; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7

DB 18 GRGDSPC 24

RESULT 12  
US-09-911-569-119  
; Sequence 119, Application US/09911569  
; Publication No. US20030069173A1  
; GENERAL INFORMATION:  
; APPLICANT: HAWLEY-NELSON, PAMELA  
; LAN, JIANQING  
; SHIH, POJEN  
; JESSE, JOEL A.  
; SCHIFFERLI, KEVIN P.  
; GEBYERHU, GULLIAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/911,569  
FILING DATE: 23-Jul-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95D  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 119:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 119;  
US-09-911-569-119

Query Match 100.0%; Score 43; DB 10; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.95; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRGDSPC 7  
DB 1 GRGDSPC 7

RESULT 13  
US-10-200-879-118  
; Sequence 118, Application US/10200879  
; Publication No. US2003014230A1  
; GENERAL INFORMATION:  
; APPLICANT: HAWLEY-NELSON, PAMELA  
; LAN, JIANQING  
; SHIH, POJEN  
; JESSE, JOEL A.

ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/200,879  
FILING DATE: 23-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/911,569  
FILING DATE: 23-JUL-2001  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
APPLICATION NUMBER: US 08/818,200  
FILING DATE: 14-MAR-1997  
APPLICATION NUMBER: US 08/658,130  
FILING DATE: 04-JUN-1996  
APPLICATION NUMBER: US 08/477,354  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 119:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 119:  
US-10-200-879-119  
Query Match 100.0%; Score 43; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
Db 1 GRGDSPC 7  
RESULT 15  
US-09-911-569-62  
Sequence 62, Application US/09911569  
Publication No. US20030069173A1  
GENERAL INFORMATION:  
APPLICANT: HAWLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBEYEHU, GULILAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US

SCHIFFERLI, KEVIN P.  
GEBEYEHU, GULILAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/200,879  
FILING DATE: 23-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/911,569  
FILING DATE: 23-JUL-2001  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
APPLICATION NUMBER: US 08/818,200  
FILING DATE: 14-MAR-1997  
APPLICATION NUMBER: US 08/658,130  
FILING DATE: 04-JUN-1996  
APPLICATION NUMBER: US 08/477,354  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 118:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 118:  
US-10-200-879-118  
Query Match 100.0%; Score 43; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GRGDSPC 7  
Db 18 GRGDSPC 24  
RESULT 14  
US-10-200-879-119  
Sequence 119, Application US/10200879  
Publication No. US2003014423A1  
GENERAL INFORMATION:  
APPLICANT: HAWLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBEYEHU, GULILAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:



us-09-991-588b-10.rapb

Fri Apr 16 11:04:08 2004

QY 1 GREGSPC 7  
 Db 70 GREGSPC 76  
 Search completed: April 16, 2004, 08:03:16  
 Job time : 47.075 secs

ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA: US/09/911,569  
 APPLICATION NUMBER: US 09/039,780  
 FILING DATE: 23-Jul-2001  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 09/039,780  
 FILING DATE: 16-MAR-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: SULLIVAN, SALLY A.  
 REGISTRATION NUMBER: 32,064  
 REFERENCE/DOCKET NUMBER: 32-95D  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303)499-8080  
 TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 62:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 96 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FEATURE:

NAME/KEY: Modified-site  
 LOCATION: 1..20  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa AT POSITIONS 1-20 CAN BE ANY AMINO ACIDS OR CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 21  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "CYS AT POSITION 21 CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 22  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa AT POSITION 22 CAN BE TYR OR TRP OR CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 23..42  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa AT POSITIONS 23-42 CAN BE ANY AMINO ACIDS OR CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 50..69  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa AT POSITIONS 50-69 CAN BE ANY AMINO ACIDS OF CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 76  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "CYS AT POSITION 76 CAN BE ABSENT"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 77..96  
 OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa AT POSITIONS 77-96 CAN BE ANY AMINO ACIDS OR CAN BE ABSENT"

SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
 US-09-911-569-62  
 Query Match 100.0%; Score 43; DB 10; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:02:47 ; Search time 31.6 Seconds  
(without alignments)

39,939 Million cell updates/sec

Title: US-09-991-588B-23

Perfect score: 21

Sequence: 1 RGDS 4

Scoring table: BLOSUM62  
Gapop -0.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL 25:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phase:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_virus:  
16: sp\_bacteriap:  
17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	20	Q9PXE4	Q9PXE4 foot-and-mo
2	21	100.0	23	Q9UC00	Q9UC00 homo sapien
3	21	100.0	24	Q9UR70	Q9UR70 aspergillus
4	21	100.0	29	Q9UDJ9	Q9UDJ9 homo sapien
5	21	100.0	43	Q9KMX9	Q9KMX9 vibrio chol
6	21	100.0	46	Q8J004	Q8J004 homo sapien
7	21	100.0	47	Q93TV9	Q93TV9 pseudomonas
8	21	100.0	47	Q9L6U0	Q9L6U0 alcaligenes
9	21	100.0	47	Q8J005	Q8J005 homo sapien
10	21	100.0	47	Q8IXB7	Q8IXB7 homo sapien
11	21	100.0	47	Q8IXB6	Q8IXB6 homo sapien
12	21	100.0	50	Q8LHC2	Q8LHC2 cryza sativ
13	21	100.0	52	Q9UDK9	Q9UDK9 homo sapien
14	21	100.0	54	Q8FMC1	Q8FMC1 corynebacte
15	21	100.0	54	Q7UVE0	Q7UVE0 rhodopirell
16	21	100.0	55	Q7JW39	Q7JW39 rhodopirell

17	21	100.0	56	4	Q9UDL0
18	21	100.0	60	16	Q83CJ4
19	21	100.0	62	10	Q8S093
20	21	100.0	64	11	Q83005
21	21	100.0	65	6	Q95M28
22	21	100.0	65	6	Q95M46
23	21	100.0	65	12	Q8V989
24	21	100.0	65	12	Q8V989
25	21	100.0	65	12	Q8V993
26	21	100.0	67	2	Q7WZ19
27	21	100.0	68	16	Q8TTP5
28	21	100.0	70	2	Q51599
29	21	100.0	70	2	Q8GHM1
30	21	100.0	70	10	Q8L9W9
31	21	100.0	70	12	Q91QS3
32	21	100.0	71	10	Q42294
33	21	100.0	72	10	Q94EX1
34	21	100.0	72	16	Q7ULT1
35	21	100.0	73	17	Q8TYS2
36	21	100.0	74	5	Q8IB34
37	21	100.0	74	8	Q98RT8
38	21	100.0	75	5	Q7YU28
39	21	100.0	76	6	Q9GME2
40	21	100.0	77	16	Q7VZS1
41	21	100.0	78	10	Q42061
42	21	100.0	80	2	Q8KSP2
43	21	100.0	83	4	Q9P012
44	21	100.0	83	16	Q87UF1
45	21	100.0	83	16	Q7WGA3

#### ALIGNMENTS

##### RESULT 1

Q9PXE4 PRELIMINARY: PRT; 20 AA.  
AC Q9PXE4;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DE 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE VP1 protein (Fragment).  
OS Foot-and-mouth disease virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=12110;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96100820; PubMed=7483796;  
RA Piatti P., Hassard S., Newman J.F., Brown F.;  
RT "Antigenic variants in a plaque-isolate of foot-and-mouth disease virus: implications for vaccine production."  
RL Vaccine 13:781-784(1995).  
SQ SEQUENCE 20 AA; 1925 MW; E91F87C99C19D7DB CRC64;

Query Match 100.0%; Score 21; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
DE 5 RGDS 8

##### RESULT 2

Q9UC00 PRELIMINARY: PRT; 23 AA.  
AC Q9UC00;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DE 01-JUN-2000 (TRENBLrel. 24, Last annotation update)  
DE Enhancement of wound HEALING process.  
OS Homo sapiens (Human).

"The human mRNA encoding the Goodpasture antigen is alternatively spliced.";  
J. Biol. Chem. 268:12090-12094 (1993).

1  
NON\_TER 18 19  
NON\_TER 29 29  
SEQUENCE 29 AA; 3102 MW; 2B7047AAB1580036 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. NO. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
|||  
DB 6 RGDS 9

RESULT 5  
ID Q9QMX9 PRELIMINARY; PRT; 43 AA.  
AC Q9QMX9  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein VCA0187.  
GN VCA0187.  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Bacteriia; Proteobacteria;  
OC Vibrionaceae; Vibrio.  
CX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=51 Tor N16961 / Serotype O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Unavay L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.A.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA Xiongnad L., Utterback T., Fleischmann R.D., Niermar W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RL Nature 406:477-483 (2000).  
DR EMBL; AE004359; RAF96100.1; --  
DR FIR; D82489; D82489.  
DR TIGR; VCA0187; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 43 AA; 4957 MW; 0077A2F778E8B40A CRC64;

Query Match 100.0%; Score 21; DB 16; Length 43;  
Best Local Similarity 100.0%; Pred. NO. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
|||  
DB 10 RGDS 13

RESULT 6  
ID Q8J004 PRELIMINARY; PRT; 46 AA.  
AC Q8J004;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Protein C (Fragment).  
GN PROC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]

RA	Vedler E., Koiv V., Heinari A.;
RT	"Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid
RT	PEST4011 of Achromobacter xylosoxidans subsp. denitrificans strain
RT	EST4002.";
DR	Gene 255;281-288(2000).
DR	EMBL; AF233095; AAF61946.1; --
DR	GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR	InterPro; IPR002513; Transposase_7.
DR	Pfam; PF01526; Transposase_7; 1.
KW	Plasmid.
FT	NON TER
SQ	SEQUENCE 47 AA; 5431 MW; 43DCEE3DEB277933 CRC64;
QY	1 RGDS 4
DB	41 RGDS 44
RESULT 3	
QB0J005	PRELIMINARY; PRT; 47 AA.
ID	QB0J005
AC	QB0J005;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
PROT	Protein C (fragment).
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	
OX	[1]
RN	SEQUENCE FROM N.A.
RP	Kinoshita S., Iida H., Inoue S., Watanabe K., Kurihara M., Wada Y.,
RA	Ono M., Dongchon K., Hamasaki N.;
RT	"Gene Analysis of Anticoagulation Factors in Japanese Thrombotic
RT	Patients. Genetic Background of Thrombophilia in Japan."
RL	Submitted (Apr-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AB083697; BAC21169.1; --	
NON TER	1
NON TER	47
SEQUENCE	47 AA; 5770 MW; 3ACAEG9E1A481932 CRC64;
Query Match	100.0%; Score 21; DB 4; Length 47;
Best Local Similarity	100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 RGDS 4
DB	41 RGDS 44
RESULT -0	
QBIXB7	PRELIMINARY; PRT; 47 AA.
ID	QBIXB7
AC	QBIXB7;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
PROT	Protein C (Fragment).
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	
OX	[1]
RN	SEQUENCE FROM N.A.
RP	Hamasaki S., Kang D., Kinoshita S., Iida K., Inoue S., Watanabe K.,
RA	Kurihara M., Wada Y., Ono M.;

RT "Gene analysis of anticoagulation factors in Japanese thrombotic  
 RT patients. Genetic background of thrombophilia in Japan.";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB086849; BAC53629.1; -  
 FT NCNTER 1 1  
 FT NONTER 47 47  
 SQ SEQUENCE 47 AA; 5770 MW; 0D7AE69E1A4F1E46 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 41 RGDS 44

## RESULT 11

Q8IXB6 PRELIMINARY; PRT; 47 AA.  
 AC Q8IXB6;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Protein C (Fragment).  
 GN PROCT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hamasaki S., Kang D., Kinoshita S., Iida K., Inoue S., Watarabe K.,  
 RA Kurihara M., Wada Y., Ono M.;  
 RT "Gene analysis of anticoagulation factors in Japanese thrombotic  
 RT patients. Genetic background of thrombophilia in Japan.";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB086850; BAC53630.1; -  
 FT NCNTER 1 1  
 FT NONTER 47 47  
 SQ SEQUENCE 47 AA; 5800 MW; 7BCAE69E1A4F1B42 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 41 RGDS 44

## RESULT 12

Q8LHC2 PRELIMINARY; PRT; 50 AA.  
 AC Q8LHC2;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE P0458E05.28 protein.  
 GN P0458E05.28.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0458E05.";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP004365; BAC05628.1; -

DR Gramene; C8LHC2; -  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IEA.  
 DR GO; GO:0008248; F:pre-mRNA splicing factor activity; IEA.  
 DR GO; GO:0006371; P:mRNA splicing; IEA.  
 DR InterPro; IPR001163; snRNP\_Sm.  
 DR Pfam; PF01423; LSM; 1.  
 SQ SEQUENCE 50 AA; 5771 MW; DB66DDAD5310AA67 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 38 RGDS 41

## RESULT 13

Q9UDK9 PRELIMINARY; PRT; 52 AA.  
 AC Q9UDK9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE GOODPASTURE antigen (Fragments).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=9527823; PubMed=7758473;  
 RA Penades J.R., Bernal D., Revert F., Johansson C., Fresquet V.J.,  
 RA Cervera J., Wieslander J., Quinones S., Saus J.;  
 RT "Characterization and expression of multiple alternatively spliced  
 RT transcripts of the Goodpasture antigen gene region. Goodpasture  
 RT antibodies recognize recombinant proteins representing the autoantigen  
 RT and one of its alternative forms.";  
 RL Eur. J. Biochem. 229:754-760(1995).  
 DR PIR; S69113; S69113.  
 FT NCNTER 1 1  
 FT NONTER 45 46  
 FT NONTER 52 52  
 SQ SEQUENCE 52 AA; 5442 MW; 046AB41B149DDAE3 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 15 RGDS 18

## RESULT 14

Q8FMC1 PRELIMINARY; PRT; 54 AA.  
 AC Q8FMC1;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN CE2586.  
 OS Corynebacterium efficiens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=152794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
 RA Kwarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,

RA Usuda Y., Sugimoto S.;  
 RT "The entire genomic sequence of *Corynebacterium efficiens* YS-314."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP005222; BAC19396.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 54 AA; 5960 MW; 60E1E37883079410 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 ||||  
 Db 5 RGDS 8

# RESULT 15

ID Q7UVE0 PRELIMINARY; PRT; 54 AA.  
 AC Q7UVE0;  
 DT 01-OCT-2003 (TREMELrel. 25, Created)  
 DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN RB2700.  
 OS Rhodopirellula baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeeling H., Lombardot T.,  
 Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 Schiesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL: BX294137; CAD72785.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 54 AA; 5781 MW; 4640CE16B141E35 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 ||||  
 Db 3 RGDS 6

Search completed: April 16, 2004, 07:27:54  
 Job time : 33.6 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:56:42 ; Search time 7 Seconds  
(without alignments)  
29.754 Million cell updates/sec

Title: US-09-991-588B-23

Perfect score: 21

Sequence: 1 RGDS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	52	1	SODD DEBHA
2	21	100.0	71	1	CSRA_XYLFA
3	21	100.0	77	1	RUXX_AERPE
4	21	100.0	89	1	MINE_YERPE
5	21	100.0	90	1	REV_HVLZH
6	21	100.0	93	1	LSM3_SCHPO
7	21	100.0	94	1	ERB3_MOUSE
8	21	100.0	109	1	PT06_STYPL
9	21	100.0	110	1	SMD2_YEAST
10	21	100.0	112	1	YN48_ARCFU
11	21	100.0	116	1	REV_HVLRH
12	21	100.0	118	1	SMD2_CABEL
13	21	100.0	118	1	SMD2_HUMAN
14	21	100.0	119	1	SMD2_DROME
15	21	100.0	121	1	UL10_HCMVA
16	21	100.0	123	1	RL17_WYCGE
17	21	100.0	123	1	Y949_METJA
18	21	100.0	135	1	DICA_ECOLI
19	21	100.0	148	1	CB2Q_XENLA
20	21	100.0	150	1	Y883_HELPJ
21	21	100.0	150	1	Y949_HELPY
22	21	100.0	151	1	RNB_HSV2H
23	21	100.0	152	1	SODC_HOMLT
24	21	100.0	153	1	SODC_ASFFU
25	21	100.0	153	1	SODC_CANAL
26	21	100.0	153	1	SODC_DEBHA
27	21	100.0	153	1	SODC_NEUCR
28	21	100.0	154	1	SODC_SCHPO
29	21	100.0	155	1	R24A_YEAST
30	21	100.0	155	1	R24B_YEAST
31	21	100.0	155	1	RL24_KLULA
32	21	100.0	157	1	SMF1_HUMAN
33	21	100.0	157	1	SMF1_MOUSE

34 21 100.0 158 1 CUI4\_HUMAN  
35 21 100.0 158 1 CUI4\_MOUSE  
36 21 100.0 159 1 DUT\_ORFN2  
37 21 100.0 164 1 RL24\_CICAR  
38 21 100.0 168 1 GRP2\_SORBI  
39 21 100.0 177 1 CRG2\_RANCA  
40 21 100.0 177 1 CRG2\_RANCA  
41 21 100.0 177 1 MSRA\_LISIN  
42 21 100.0 177 1 MSRA\_LISMO  
43 21 100.0 178 1 THF1\_ARATH  
44 21 100.0 182 1 THF1\_BRANA  
45 21 100.0 184 1 RK5\_MESVI

P56557 homo sapien  
Q9DIX9 mus musculu  
E14597 ori virus (  
O65743 cicer ariet  
Q99070 sorghum bic  
Q91320 rana catesb  
Q91321 rana catesb  
Q92ae8 listeria in  
Q9y640 listeria mo  
Q9xfh8 arabidopsis  
O48897 brassica na  
Q9muu5 mesostigma

#### ALIGNMENTS

RESULT 1  
SODD\_DEBHA STANDARD; PRT; 52 AA.  
AC P82902;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-MAR-2004 (Rel. 43, Last annotation update)  
DE Superoxide dismutase [Cu-Zn] 2 (EC 1.15.1.1) (Fragments).  
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.  
OX NCBI\_TaxID=4959;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CIBNOR C-11;  
RA Hernandez-Saavedra N.Y.;  
RT "Presence of two active forms of cytosolic Cu-Zn superoxide dismutase enzyme in the marine yeast Debaryomyces hansenii.";  
RL Submitted (DEC-2000) to Swiss-Prot  
CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H2O(2).  
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By similarity).  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR PROSITE; PS00087; SOD\_CU\_ZN\_1; PARTIAL.  
DR PROSITE; PS00332; SOD\_CU\_ZN\_2; PARTIAL.  
KW Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc.  
FT NON\_TER 1 1  
FT NON\_CONS 29 30 COPPER (BY SIMILARITY).  
FT METAL 44 44  
FT NON\_TER 52 52  
SQ SEQUENCE 52 AA; 5555 MW; 830DA15280D25708 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 52;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4

DB 7 RGDS 10

#### RESULT 2

CSRA\_XYLFA STANDARD; PRT; 71 AA.  
AC Q9PHZ1; Q87F42;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Carbon storage regulator homolog.  
GN CSRA OR XF0125 OR PD0095.  
OS Xylella fastidiosa, and

OS Xylella fastidiosa (strain Temeculal / ATCC 700964).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xylella.

OX NCBI\_TaxID=2371, 183190;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9a5C;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reirach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.X., Carrier H.,

RA Colauto M.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferris J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,

RA Garnier Y., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA He P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.E., Kuranae E.F., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.S.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.P., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Resquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira M.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,

RA Vallada H., van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-159(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Temeculal / ATCC 700964;

RX MEDLINE=22421331; PubMed=12533478;

RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,

RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,

RA Takita M.A., Lemos E.G.M., Machado M.A., Perro M.I.T., da Silva F.R.,

RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,

RA Carrier H., Carrao D.M., de Oliveira R.C., Nunes L.R., Siqueira M.J.,

RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,

RA Marino C.L., Gagliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,

RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,

RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,

RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,

RA de Souza A.A., Truffi D., Tsubako M.H., Yanai G.M., Zago L.G.,

RA Cliverolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,

RA Kitajima J.P.;

RT "Comparative analyses of the complete genome sequences of Pierce's

RT disease and citrus variegated chlorosis strains of Xylella

RT fastidiosa.";

RL J. Bacteriol. 185:1018-1026(2003).

CC -!- FUNCTION: Could accelerate the degradation of some genes

CC transcripts potentially through selective RNA binding (by

CC similarity).

CC -!- SIMILARITY: Belongs to the crfA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AE003866; AAF82938.1; ALT -INIT.

DR EMBL; AE012553; AAC27995.1; -.

DR PIR; D82844; D82844.

DR

DR HAMAP; MF 00167; -; 1.

DR InterPro; IPR003751; CsrA.

DR Pfam; PF02599; CsrA; 1.

DR ProDom; PD009007; CsrA; 1.

DR TIGRFAMs; TIGR00202; csrA; 1.

KW RNA-binding; Complete proteome.

FT DOMAIN 11 45 KH.

SQ SEQUENCE 71 AA; 7672 MW; FABA182A40CE72AD CRC64;

Query Match 100.0%; Score 21; DB 1; Length 71;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4

Db 64 RGDS 67

RESULT 3

RUXX AERPE

ID RUXX\_AERPE STANDARD; PRT; 77 AA.

AC Q9YEQ5;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)

DE Putative snRNP Sm-like protein.

GN APES022.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;

OC Desulfurococaceae; Aeropyrum.

OX NCBI\_TaxID=56636;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KJ;

RX MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,

RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RT "Complete genome sequence of an aerobic hyper-thermophilic

RT crenarchaeon, Aeropyrum pernix K1.";

RL DNA Res. 6:83-101(1999).

CC -!- SIMILARITY: Belongs to the snRNP Sm proteins family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AP000059; BAA79491.1; -.

DR PIR; G72749; G72749.

DR HGSP; P14678; LD3B.

DR HAMAP; MF\_00257; -; 1.

DR InterPro; IPR001163; snRNP Sm.

DR Pfam; PF01423; LSM; 1.

DR SMART; SM00651; Sm; 1.

KW Hypothetical protein; Ribonucleoprotein; Complete proteome.

SQ SEQUENCE 77 AA; 8374 MW; 41CC5514366B453C CRC64;

Query Match 100.0%; Score 21; DB 1; Length 77;

Best Local Similarity 100.0%; Pred. No. 73;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4

Db 66 RGDS 69



```

RESULT 4
MINE_YERPE
ID MINE_YERPE STANDARD; PRT; 89 AA.
AC Q82ES7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell division topological specificity factor.
GN MINE OR YPO2076 OR Y2234.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wret B.M., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Meule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Soutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
CC -!- FUNCTION: Prevents the cell division inhibition by proteins minC
CC and minD at internal division sites while permitting inhibition at
CC polar sites. This ensures cell division at the proper site by
CC restricting the formation of a division septum at the midpoint of
CC the long axis of the cell (by similarity).
CC -!- SIMILARITY: Belongs to the minE family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ414151; CAC90888.1; -.
CC DR EMBL; AE013826; AAM85794.1; -.
CC DR PIR; A02053; AD0253.
CC DR HAMAP; MF_00262; -.
CC DR InterPro; IPR005527; MinE.
CC DR Pfam; PF03776; MinE; 1.
CC DR TIGRFAMs; TIGR01215; minE; 1.
CC KW Cell division; Complete proteome.
SQ SEQUENCE 89 AA; 10332 MW; 960E226BAAF0B27B CRC64;
Query Match 100.0%; Score 21; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDS 4
| | | |
DB 31 RGDS 34
RESULT 5
REV_HV1ZH
ID REV_HV1ZH STANDARD; PRT; 90 AA.
AC P05868;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS)
DE (Fragment).
GN REV.
OS Human immunodeficiency virus type 1 (Zaire H321 isolate) (HIV-1).
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; PubMed=2713163;
RA Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchall J.,
RA McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;
RT "Molecular characterization of HIV-1 isolated from a serum collected
RT in 1976: nucleotide sequence comparison to recent isolates and
RT generation of hybrid HIV."
RL AIDS Res. Hum. Retroviruses 5:121-129(1989).
CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC -!- NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -!- PTM: Phosphoprotein whose state of phosphorylation is mediated by
CC a specific serine kinase activity present in the nucleus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M15896; AAB53950.1; -.
CC DR HIV; M15896; REV$2321.
CC DR InterPro; IPR000625; REV_protein.
CC DR Pfam; PF00424; REV; 1.
CC KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 90 AA; 10173 MW; 32F4D58657D7DB83 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDS 4
| | | |
DB 63 RGDS 66
RESULT 6
LSM3 SCHPO
ID LSM3 SCHPO STANDARD; PRT; 93 AA.
AC Q9Y7M4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM3.
GN SPBC986.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne X., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

```

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA McNeely P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehmach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cervitti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Jessery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of *Schizosaccharomyces pombe*.";  
RL Nature 4.5:871-880(2002).  
CC -!- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6  
CC snRNA (By similarity).  
CC -!- SUBUNIT: LSM subunits form a heteromer with a doughnut shape (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- SIMILARITY: Belongs to the snRNP Sm proteins family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AL049763; CAB42366.2; -.  
CC GenBank: Sponbe; SPBC3B6.05C; -.  
CC InterPro: IPR001163; snRNP\_Sm.  
CC Pfam: PF01423; LSM; 1.  
CC SMART: SM00651; Sm; 1.  
CC KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;  
CC RNA-binding.  
CC SQ SEQUENCE 93 AA; 10665 MW; 86874A78670B1DCE CRC64;  
Query Match 100.0%; Score 21; DB 1; Length 93;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDS 4  
Db 81 RGDS 84  
|||||  
RESULT 7  
ERB3\_MOUSE STANDARD; PRT; 94 AA.  
AC Q61526;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-3 (EC 2.7.1.112) (c-erbB3  
DE {Fragment}).  
GN ERB3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96069911; PubMed=7589796;  
RA Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,

RA Sanes J.R.;  
RT "Synapse-associated expression of an acetylcholine receptor-inducing  
RT protein, ARIA/heretulin, and its putative receptors, ErbB2 and ErbB3,  
RT in developing mammalian muscle.";  
RL Dev. Biol. 172:158-169(1995).  
CC -!- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTA (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
CC (POTENTIAL).  
CC -!- SUBCELLULAR LOCATION: TYPE I membrane protein (By similarity).  
CC -!- TISSUE SPECIFICITY: IN THE MUSCLE, EXPRESSION LOCALIZES TO THE  
CC SYNAPTIC SITES OF MUSCLE FIBERS.  
CC -!- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE  
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.  
CC -!- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES  
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF  
CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).  
CC -!- SIMILARITY: Belongs to the EGFR receptor family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L47240; AAA93533.1; -.  
CC M3D: MGI-95411; ErbB3.  
CC GO: GO:0007507; P:heart development; IMP.  
CC GO: GO:0007422; P:peripheral nervous system development; IMP.  
CC InterPro: IPR000719; Prot Kinase.  
CC InterPro: IPR008266; Tyr kinase AS.  
CC PROSITE: PS00107; PROTEIN KINASE ATP; PARTIAL.  
CC PROSITE: PS00109; PROTEIN KINASE TYR; PARTIAL.  
CC PROSITE: PS50011; PROTEIN KINASE DOM; PARTIAL.  
CC Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;  
CC Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
CC FT NON TER 1 94  
CC FT NON TER 94  
CC SQ SEQUENCE 94 AA; 10320 MW; 49CE4DFE49191AAB CRC64;  
Query Match 100.0%; Score 21; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDS 4  
Db 64 RGDS 67  
|||||  
RESULT 8  
PT06\_STYPL STANDARD; PRT; 109 AA.  
AC P28198;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Protein-tyrosine phosphatase 6 (EC 3.1.3.48) {Fragment}.  
GN STY 6.  
OS Styela plicata (Sea squirt).  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
OC Stolidobranchia; Styelidae; Styela.  
OX NCBI\_TaxID=7726;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91139172; PubMed=1704870;  
RA Matthews R.O., Flores E., Thomas M.L.;  
RT "Protein tyrosine phosphatase domains from the protochordate *Styela*  
RT *plicata*.";  
RL Immunogenetics 33:33-41(1991).

CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein  
 CC tyrosine + phosphate.  
 CC -!- MISCELLANEOUS: STY 3 and STY 6 may represent a double domain  
 CC transmembrane protein tyrosine phosphatase.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M37991; AAA29824.1; -;  
 DR HSPB; P28827; IREB.  
 DR InterPro; IPR000387; TYR phosphatase.  
 DR InterPro; IPR000242; TYR PP.  
 DR Pfam; PF0102; Y-phosphatase; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; PARTIAL.  
 DR PROSITE; PS0056; TYR\_PHOSPHATASE\_2; PARTIAL.  
 DR PROSITE; PS0055; TYR\_PHOSPHATASE\_FTP; 1.  
 KW Hydrolase.  
 FT NON TER 1 1  
 FT NON TER 109 109  
 SQ SEQUENCE 109 AA; 12744 MW; BED2D6F524AE88EB CRC64;

Query Match 100.0%; Score 21; DB 1; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 27 RGDS 30

RESULT 9  
 SMD2 YEAST  
 ID -SMD2 YEAST STANDARD; PRT; 110 AA.  
 AC Q06217;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2003 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Small nuclear ribonucleoprotein Sm D2 (snRNP core protein D2) (Sm-D2).  
 GN SMD2 OR YLR275W OR L9328.5.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 R2 SEQUENCE FROM N.A.  
 R3 MEDLINE=98187616; PubMed=9528767;  
 RA Camases A., Bragado-Nilsson E., Martin R., Seraphin B., Bordonne R.;  
 RT "Interactions within the yeast Sm core complex: from proteins to amino  
 RT acids.";  
 RL Mol. Cell. Biol. 18:1956-1966(1998).  
 RN [2]  
 R2 SEQUENCE FROM N.A., AND RNA-BINDING.  
 R3 MEDLINE=9298196; PubMed=10369684;  
 RA Salgado-Garrido J., Bragado-Nilsson E., Kandel-Lewis S., Seraphin B.;  
 RT "Sm and Sm-like proteins assemble in two related complexes of deep  
 RT evolutionary origin.";  
 RL EMBO J. 18:3451-3462(1999).  
 RN [3]  
 R2 SEQUENCE FROM N.A.  
 R3 STRAIN=S288C / AB972;  
 MEDLINE=97313267; PubMed=9169871;  
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,  
 RA Benes V., Brueckner M., Delius H., Dubois E., Dueterhoeft A.,  
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,  
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,  
 RA Louis E.J., Messenguy F., Mewes H.-W., Miesga T., Moestl D.,  
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,  
 RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,

RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,  
 RA Underwood A.P., Errestazu L.A., Vandenbol M., Verhasselt P.,  
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,  
 RA Wiering H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";  
 RL Nature 387:87-90(1997).  
 RN [4]  
 RP REVISIONS.  
 RA Cherry J.M.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBSJ databases.  
 RN [5]  
 RP PARTIAL SEQUENCE, AND IDENTIFICATION.  
 RX MEDLINE=97165029; PubMed=9012791;  
 RA Neubauer G., Gottschalk A., Fabrizio P., Seraphin B., Luehrmann R.,  
 RA Mann M.;  
 RT "Identification of the proteins of the yeast U1 small nuclear  
 RT ribonucleoprotein complex by mass spectrometry.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:385-390(1997).  
 RN [6]  
 RP CHARACTERIZATION OF THE SPLICEOSOME.  
 RX MEDLINE=21664121; PubMed=11804584;  
 RA Stevens S.W., Ryan D.E., Ge H.Y., Moore R.E., Young M.K., Lee T.D.,  
 RA Abelson J.;  
 RT "Composition and functional characterization of the yeast spliceosomal  
 RT pentamer snRNP.";  
 RL Mol. Cell 9:31-44(2002).  
 RN [7]  
 RP SUBUNIT.  
 RX MEDLINE=21199560; PubMed=11302706;  
 RA Walke S., Bragado-Nilsson E., Seraphin B., Nagai K.;  
 RT "Stoichiometry of the Sm proteins in yeast spliceosomal snRNPs  
 RT supports the heptamer ring model of the core domain.";  
 RL J. Mol. Biol. 308:49-58(2001).  
 CC -!- FUNCTION: Involved in pre-mRNA splicing. Binds sn-RNA U1, U2, U4  
 CC and U5 which contain a highly conserved structural motif called  
 CC the Sm binding site.  
 CC -!- SUBUNIT: Component of the Sm core complex, present in spliceosomal  
 CC snRNP U1, U2, U4/U6 and U5. The core complex contains SYB1, SMD1,  
 CC SMD2, SMD3, SMEL, SMX3 and SMX2 (Sm proteins B, D1, D2, D3, E, F  
 CC and G, respectively), and is probably a heptameric ring structure.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Belongs to the snRNP core protein family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U17245; AAB67368.1; -;  
 DR PIR; S69326; S69326.  
 DR GeneOnline; 142337; -;  
 DR SGD; S0004265; SMD2.  
 DR GO; GO:0030532; C:small nuclear ribonucleoprotein complex; IDA.  
 DR GO; GO:0008248; F:pre-mRNA splicing factor activity; IPI.  
 DR GO; GO:0006371; P:mRNA splicing; IPI.  
 DR InterPro; IPR001163; snRNP\_Sm.  
 DR Pfam; PF01423; LSM; 1.  
 DR SMART; SM00651; Sm; 1.  
 KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;  
 KW RNA-binding.  
 SQ SEQUENCE 110 AA; 12853 MW; 7287070B10E090C0 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 DB 97 RGDS 100

```
RESULT 10
YN48_ARCFU STANDARD; PRT; 112 AA.
AC C030321;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein AF2348.
GN AF2348.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Karlavage A.R., Graham D.B., Kyripides N.C., Gill S.,
RA Fleischmann R.D., Gachekush J., Lee N.H., Sutton G.S., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.W., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001114; AAB91315.1; -.
CC DR PIR; D69543; D69543.
CC DR TIGR; AF2348; -.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 112 AA; 13012 MW; C9C6ACACD6AC730 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGS 4
DB 16 REGS 19

RESULT 11
REV_HVIRH STANDARD; PRT; 116 AA.
AC P05870;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/IRF).
GN REV.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RA Starcich B.R., Rahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;

Submitted (XX-1987) to the HIV data bank.
-!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
-!- NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
-!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
-!- PTM: Phosphoprotein whose state of phosphorylation is mediated by
a specific serine kinase activity present in the nucleus.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17451; AAA45051.1; -.
CC DR HIV; M17451; REVSRF.
CC DR InterPro; IPR000625; REV_protein.
CC PFam; PF03424; REV; 1.
CC KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
CC SQ SEQUENCE 116 AA; 13113 MW; 5CC86F838F40CD12 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGS 4
DB 5 REGS 8

RESULT 12
SMD2_CABEL STANDARD; PRT; 118 AA.
AC Q-8786;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable small nuclear ribonucleoprotein Sm D2 (snRNP core protein D2)
DE (Sm-D2).
GN SNR-4 OR C52E4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Holt R.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Required for pre-mRNA splicing. Required for snRNP
CC biogenesis (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the snRNP core protein family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z78012; CAB01413.1; -.
CC DR PIR; Z78012; T20151.
CC DR WormPep; C52E4.3; CE08945.
CC DR InterPro; IPR01163; snRNP_Sm.
CC PFam; PF01423; LSM; 1.
CC SMART; SM00651; Sm; 1.
CC KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing.
CC SQ SEQUENCE 118 AA; 13267 MW; 24773D27A1D1555 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 118;
```

Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
Db 102 RGDS 105

RESULT 13  
SMD2\_HUMAN STANDARD; PRT; 118 AA.  
AC P43330;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Small nuclear ribonucleoprotein Sm D2 (snRNP core protein D2) (Sm-D2).  
GN SNRPD2.  
OS Homo sapiens (Human), and  
OS Mus musculus (Mouse).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606, 10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 75-96 AND 103-118.  
RC SPECIES=Human;  
RX MEDLINE=95083692; PubMed=7527560;  
RA Lehwieser T., Raker V., Hermann H., Luehrmann R.;  
RT "cDNA cloning of the Sm proteins D2 and D3 from human small nuclear  
ribonucleoproteins: evidence for a direct D1-D2 interaction.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:12317-12321(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
RA Phan H., Velasco N., Do L., Regala M., Terry A., Ganes J.,  
RA Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carraro A.V.;  
RT "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and  
D19S412.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feilgenfeldt E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzinski W.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse; STRAIN=C57BL/6J;  
RX TISSUE=Corpora quadrigemina, Pancreas, and Thymus;  
MEDLINE=21085660; PubMed=11217851;

Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,  
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
Nordone P., Ring B., Kingwald M., Rodriguez I., Sakamoto N.,  
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse; STRAIN=C57BL/6; TISSUE=Brain;  
RX Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Required for pre-mRNA splicing. Required for snRNP  
CC -!- SUBCELLULAR LOCATION: Nucleus.  
CC -!- DISEASE: In the autoimmune disease systemic lupus erythematosus,  
CC antinuclear antibodies are developed with Sm specificity.  
CC -!- SIMILARITY: Belongs to the snRNP core protein family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U15008; AAC13776.1; -;  
CC EMBL; AC007191; AAD22673.1; -;  
CC EMBL; BC000486; AAH00486.1; -;  
CC EMBL; BC001930; AAH01930.1; -;  
CC EMBL; AK007389; BAB25006.1; -;  
CC EMBL; AK045968; BAC32551.1; -;  
CC EMBL; AK088105; BAC40147.1; -;  
CC EMBL; BC043014; AAH43014.1; -;  
CC PIR; I38861; I38861.  
CC PDB; 1B34; 1J-JAN-00  
CC Genew; HGNC:11159; SNRPD2.  
CC GK; P43330; -;  
CC MED; MGI:98345; Snrpd2.  
CC GO; GO:0030532; C:small nuclear ribonucleoprotein complex; TAS.  
CC GO; GO:0005681; C:spliceosome complex; TAS.  
CC GO; GO:0006371; P:mRNA splicing; TAS.  
CC GO; GO:0000245; P:spliceosome assembly; TAS.  
CC InterPro; IPR001163; snRNP\_Sm.  
CC Pfam; PF01423; LSM; 1.  
CC SMART; SM00651; Sm; 1.  
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;  
KW Systemic lupus erythematosus; 3D-structure.  
FT CONFLICT 31 31 V -> M [IN REF. 4; BAC32551].  
SQ SEQUENCE 118 AA; 13527 MW; D986059D82B7E045 CRC64;  
Query Match 100.0%; Score 21; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
Db 102 RGDS 105

```

RESULT 14
SMD2_DROME
ID_SMD2_DROME STANDARD; PRT: 119 AA.
AC QSV10;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Probable small nuclear ribonucleoprotein Sm D2 (snRNP core protein D2)
DE (Sm-D2).
DE CGL249.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Baze J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale C., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman G.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cayley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimme B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.X.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RN Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.M., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celnik S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -!- FUNCTION: Required for pre-mRNA splicing. Required for snRNP
CC biogenesis (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the snRNP core protein family.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE003675; AAF54135.1; -  
EMBL; AY089315; AAL39460.1; -  
FlyBase; F5gn0037434; CG1249.  
DR GO; GO:0030532; C:small nuclear ribonucleoprotein complex; ISS.  
DR GO; GO:0008248; P:pre-mRNA splicing factor activity; ISS.  
DR GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; ISS.  
DR InterPro; IPR001163; snRNP\_Sm.  
DR Pfam; PF01423; LSM; 1.  
DR SMART; SMO0651; Sm; 1.  
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing.  
SQ SEQUENCE 119 AA; 13504 MW; F292C54E400093B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
DB 101 RGDS 104

RESULT 15  
UL30\_HCMVA STANDARD; PRT: 121 AA.  
ID UL30\_HCMVA STANDARD; PRT: 121 AA.  
AC P16765;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Hypothetical protein UL30.  
GN UL30.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
RA Horsnell T., Hutchinson C.A. III, Kouravides T., Martignetti J.A.,  
RA Reddick E., Satchwell S.C., Tomlinson P., Weston K.M., Barrett B.G.;  
RT "Analysis of the protein-coding content of the sequence of human  
RT cytomegalovirus strain AD169".  
RL Curr. Top. Microbiol. Immunol. 154:125-169 (1990).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X17403; CAA35429.1; -  
PIR; S09793; S09793.  
KW Hypothetical protein.  
SQ SEQUENCE 121 AA; 14046 MW; 52DC9C015DC5BC54 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
DB 26 RGDS 29

Search completed: April 16, 2004, 07:22:17  
Job time : 8 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: April 16, 2004, 07:28:08 ; Search time 26.9 Seconds

(without alignments)  
40.992 Million cell updates/sec

Title: US-09-991-588b-23

Perfect score: 21

Sequence: 1 RGDS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	9	US-09-991-588b-23
2	21	100.0	4	9	US-09-991-588b-23
3	21	100.0	4	9	US-09-991-588b-23
4	21	100.0	4	9	US-09-991-588b-23
5	21	100.0	4	9	US-09-991-588b-23
6	21	100.0	4	10	US-09-991-588b-23
7	21	100.0	4	11	US-09-991-588b-23
8	21	100.0	4	12	US-09-991-588b-23
9	21	100.0	4	12	US-09-991-588b-23
10	21	100.0	4	14	US-09-991-588b-23
11	21	100.0	4	14	US-09-991-588b-23
12	21	100.0	4	14	US-09-991-588b-23
13	21	100.0	4	14	US-09-991-588b-23
14	21	100.0	4	14	US-09-991-588b-23
15	21	100.0	4	14	US-09-991-588b-23

16	21	100.0	4	14	US-09-991-588b-23
17	21	100.0	4	14	US-09-991-588b-23
18	21	100.0	4	15	US-09-991-588b-23
19	21	100.0	4	15	US-09-991-588b-23
20	21	100.0	4	16	US-09-991-588b-23
21	21	100.0	5	9	US-09-991-588b-23
22	21	100.0	5	9	US-09-991-588b-23
23	21	100.0	5	10	US-09-991-588b-23
24	21	100.0	5	10	US-09-991-588b-23
25	21	100.0	5	11	US-09-991-588b-23
26	21	100.0	5	11	US-09-991-588b-23
27	21	100.0	5	12	US-09-991-588b-23
28	21	100.0	5	12	US-09-991-588b-23
29	21	100.0	5	13	US-09-991-588b-23
30	21	100.0	5	14	US-09-991-588b-23
31	21	100.0	5	14	US-09-991-588b-23
32	21	100.0	5	14	US-09-991-588b-23
33	21	100.0	5	14	US-09-991-588b-23
34	21	100.0	5	14	US-09-991-588b-23
35	21	100.0	5	14	US-09-991-588b-23
36	21	100.0	5	14	US-09-991-588b-23
37	21	100.0	5	14	US-09-991-588b-23
38	21	100.0	5	14	US-09-991-588b-23
39	21	100.0	5	14	US-09-991-588b-23
40	21	100.0	5	16	US-09-991-588b-23
41	21	100.0	6	9	US-09-991-588b-23
42	21	100.0	6	9	US-09-991-588b-23
43	21	100.0	6	9	US-09-991-588b-23
44	21	100.0	6	9	US-09-991-588b-23
45	21	100.0	6	9	US-09-991-588b-23

## ALIGNMENTS

RESULT 1  
US-09-823-444-6  
Sequence 6, Application US/09823444  
Patent No. US20020039753A1  
GENERAL INFORMATION:  
APPLICANT: Bednar, Bohumil  
APPLICANT: Bolla, Daniel M.  
APPLICANT: Gould, Robert J.  
APPLICANT: Merck & Co., Inc.  
TITLE OF INVENTION: ANTICOAGULANT TEST  
FILE REFERENCE: 19910  
CURRENT APPLICATION NUMBER: US/09/823,444  
CURRENT FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 4  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: This sequence is a synthetically prepared peptide.  
US-09-823-444-6

Query Match 100.0%; Score 21; DB 9; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDS 4  
DB 1 RGDS 4

RESULT 2  
US-09-010-714-9  
Sequence 9, Application US/09010714  
Patent No. US20020012942A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, James B.



```
; APPLICANT: Furcht, Leo T.
; APPLICANT: Iida, Joji
; TITLE OF INVENTION: POLYPEPTIDES WITH ALPHA 4 INTEGRIN SUBUNIT RELATED
; FILE REFERENCE: 60C.332US01
; CURRENT APPLICATION NUMBER: US/09/010,714
; CURRENT FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 9
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-010-714-9

Query Match      100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 3
US-09-925-715-21
; Sequence 21, Application US/09925715
; Patent No. US20020102217A1
; GENERAL INFORMATION:
; APPLICANT: Nycomed Imaging AS
; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic
; FILE REFERENCE: REF/Klaveness/206
; CURRENT APPLICATION NUMBER: US/09/925,715
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic RGDS
; OTHER INFORMATION: sequence
US-09-925-715-21

Query Match      100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 4
US-09-935-168-1
; Sequence 1, Application US/09935168
; Patent No. US20020106793A1
; GENERAL INFORMATION:
; APPLICANT: West, Jennifer L.
; APPLICANT: Mann, Brenda K.
; TITLE OF INVENTION: Tissue Engineering Scaffolds Promoting Matrix Protein Production
; FILE REFERENCE: RICE 103
; CURRENT APPLICATION NUMBER: US/09/935,168
; CURRENT FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
```

```
; OTHER INFORMATION: cell adhesion ligand
US-09-935-168-1

Query Match      100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 5
US-09-942-117-5
; Sequence 5, Application US/09942117
; Publication No. US20020197700A1
; GENERAL INFORMATION:
; APPLICANT: MENEAD, ANDREAS
; APPLICANT: REDLITZ, ALEXANDER
; APPLICANT: KOPPLITZ, MARCUS
; APPLICANT: EGNER, URSULA
; APPLICANT: BAHR, INKE
; TITLE OF INVENTION: RECEPTOR OF THE EDB-FIBRONECTIN DOMAINS
; FILE REFERENCE: SCH-1832
; CURRENT APPLICATION NUMBER: US/09/342,117
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: DE 1004S803.3
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: DE 10123133.4-42
; PRIOR FILING DATE: 2001-05-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-942-117-5

Query Match      100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 6
US-09-911-569-22
; Sequence 22, Application US/09911569
; Publication No. US20030069173A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NELSON, PAMELA
; APPLICANT: LAN, JIANQING
; APPLICANT: SHIH, POJEN
; APPLICANT: JESSE, JOEL A.
; APPLICANT: SCHIFFERLI, KEVIN P.
; APPLICANT: GEBREYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESS: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DCS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/911,569
; FILING DATE: 23-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-911-569-22

```

```

Query Match      100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RGDS 4
Db      1 RGDS 4

```

## RESULT 7

```

US-09-991-588B-23
; Sequence 23, Application US/09991588B
; Publication No. US20030219429A1
; GENERAL INFORMATION:
; APPLICANT: Badny, John A.
; TITLE OF INVENTION: Compositionand Method for Bone Regeneration
; FILE REFERENCE: 1008-120.US
; CURRENT APPLICATION NUMBER: US/09/991,588B
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 09/122,348
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 23
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Purchased commercially or sequence is synthesized
US-09-991-588B-23

```

```

Query Match      100.0%; Score 21; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RGDS 4
Db      1 RGDS 4

```

## RESULT 8

```

US-10-169-085-23
; Sequence 23, Application US/10169085
; Publication No. US20040033215A1
; GENERAL INFORMATION:
; APPLICANT: Kobayashi, Naoya

```

```

; APPLICANT: Philippe, Leboulch
; APPLICANT: Tanaka, No. US20040033215A1iaki
; APPLICANT: Fujiwara, Toshiyoshi
; APPLICANT: Toshiyori, Totugawa
; TITLE OF INVENTION: METHOD FOR PROLIFERATING A LIVER CELL, A LIVER CELL OBTAINED
; TITLE OF INVENTION: THEREBY, AND USE THEREOF
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/169,085
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/809,187
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 23
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell adhesion peptide
US-10-169-085-23

```

```

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RGDS 4
Db      1 RGDS 4

```

## RESULT 9

```

US-09-961-381A-1
; Sequence 1, Application US/09961381A
; Publication No. US20020061515A1
; GENERAL INFORMATION:
; APPLICANT: Lynch, Gary
; APPLICANT: Bi, Xiaoning
; APPLICANT: Gall, Christine M.
; TITLE OF INVENTION: Model for Neurodegenerative Diseases Involving Amyloid
; TITLE OF INVENTION: Accumulation
; FILE REFERENCE: 1819.0040001
; CURRENT APPLICATION NUMBER: US/09/961,381A
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,374
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-961-381A-1

```

```

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RGDS 4
Db      1 RGDS 4

```

## RESULT 10

```

US-10-046-801-4
; Sequence 4, Application US/10046801
; Publication No. US20030054027A1
; GENERAL INFORMATION:
; APPLICANT: Unger, Evan C.
; TITLE OF INVENTION: Charged Lipids and Uses For The Same
; FILE REFERENCE: UNGRI592
; CURRENT APPLICATION NUMBER: US/10/046,801
; CURRENT FILING DATE: 2002-05-13

```

```
; PRIOR APPLICATION NUMBER: US/09/540,448
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 08/925,353
; PRIOR FILING DATE: 1997-03-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: No. US20030054027A1el Sequenc
US-10-046-801-4

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 11
US-10-215-435-5
; Sequence 5, Application US/10215435
; Publication No. US20030104614A1
; GENERAL INFORMATION:
; APPLICANT: Utrich, Kathryn E.
; APPLICANT: Buettner, Helen
; APPLICANT: Schmalenberg, Kristine
; TITLE OF INVENTION: Micropatterning Surfaces of Polymeric Substrates
; FILE REFERENCE: 1435.0080S1
; CURRENT APPLICATION NUMBER: US/10/215,435
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: PCT/US01/04842
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: US 60/181,763
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: A peptide sequence found in fibronectin.
US-10-215-435-5

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 12
US-10-200-879-22
; Sequence 22, Application US/10200879
; Publication No. US20030144230A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NELSON, PAMELA
; APPLICANT: SHIH, JIANQING
; APPLICANT: SHIH, POJEN
; APPLICANT: JESSE, JOEL A.
; APPLICANT: SCHIFFERLI, KEVIN P.
; APPLICANT: GEBEYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
```

```
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/200,879
; FILING DATE: 23-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/911,569
; FILING DATE: 23-JUL-2001
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; APPLICATION NUMBER: US 08/818,200
; FILING DATE: 14-MAR-1997
; APPLICATION NUMBER: US 08/658,130
; FILING DATE: 04-JUN-1996
; APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-200-879-22

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RGDS 4
Db      1 RGDS 4

RESULT 13
US-10-021-660-130
; Sequence 130, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 130
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: adhesive motif
US-10-021-660-130

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. NO. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDS 4
      ||||
DB      1 RGDS 4

RESULT 14
US-10-299-043-1
; Sequence 1, Application US/10299043
; Publication NO. US20030158115A1
; GENERAL INFORMATION:
; APPLICANT: TOBACK, F. GARY
; APPLICANT: LIESKE, JOHN C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING AND TREATING
; TITLE OF INVENTION: KIDNEY DISEASES ASSOCIATED WITH ADHESION OF CRYSTALS TO
; TITLE OF INVENTION: KIDNEY CELLS
; FILE REFERENCE: 21459/90606
; CURRENT APPLICATION NUMBER: US/10/299,043
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US/09/537,226
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 08/389,005
; PRIOR FILING DATE: 1995-02-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-299-043-1

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. NO. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDS 4
      ||||
DB      1 RGDS 4

RESULT 15
US-10-300-694A-96
; Sequence 96, Application US/10300694A
; Publication NO. US20030185870A1
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Grinstaff, Mark W.
; APPLICANT: Kenan, Daniel J.
; APPLICANT: Walsh, Elisabeth B.
; APPLICANT: Middleton, Crystan
; TITLE OF INVENTION: INTERFACIAL BIOMATERIALS
; FILE REFERENCE: 180/143/2
; CURRENT APPLICATION NUMBER: US/10/300,694A
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/331,943
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 96
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cell binding peptide 96
US-10-300-694A-96

Query Match      100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. NO. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGDS 4
      ||||
DB      1 RGDS 4

Search completed: April 16, 2004, 08:03:17
Job time : 27.9 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:03:27 ; Search time 8.4 Seconds  
(without alignments)  
45.805 Million cell updates/sec

Title: US-09-991-588B-23

Perfect score: 1 RGDS 4

Sequence: 1 RGDS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: Piri.\*  
2: Piri.\*  
3: Piri.\*  
4: Piri.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	20	2 A44773	pollen allergen I
2	21	100.0	30	2 S15982	arietin - puff add
3	21	100.0	43	2 D82489	hypothetical prote
4	21	100.0	52	2 S69113	collagen alpha 3(I
5	21	100.0	53	2 S03486	T-cell receptor be
6	21	100.0	62	2 S18571	otcz protein - Str
7	21	100.0	68	2 AF2139	hypothetical prote
8	21	100.0	74	2 C90091	small nucleolar ri
9	21	100.0	76	2 D82844	carbon storage reg
10	21	100.0	77	2 G72749	probable small nuc
11	21	100.0	79	2 S17389	T-cell receptor be
12	21	100.0	81	2 A49736	collagen alpha 3(I
13	21	100.0	84	2 AC1086	hypothetical prote
14	21	100.0	87	2 JQ1683	UL49A protein - hu
15	21	100.0	88	2 T14944	hypothetical prote
16	21	100.0	89	2 AD0253	cell division topo
17	21	100.0	92	2 JH0716	neuropeptide Y pre
18	21	100.0	100	2 AI0771	hypothetical prote
19	21	100.0	101	2 D71604	small nuclear ribo
20	21	100.0	101	2 G72450	hypothetical prote
21	21	100.0	102	2 D72518	hypothetical prote
22	21	100.0	103	2 D70690	hypothetical prote
23	21	100.0	105	2 T45398	hypothetical prote
24	21	100.0	108	2 T00420	probable small nuc
25	21	100.0	110	2 S22896	T-cell receptor al
26	21	100.0	110	2 S23368	T-cell receptor al
27	21	100.0	110	2 S69326	small nuclear ribo
28	21	100.0	111	2 T49579	hypothetical prote
29	21	100.0	112	2 D69543	hypothetical prote

30	21	100.0	114	2	PT0733	T-cell receptor be
31	21	100.0	115	2	S22035	T-cell receptor be
32	21	100.0	116	2	PH1540	Ig H chain V regio
33	21	100.0	116	2	C83492	hypothetical prote
34	21	100.0	118	2	T20151	hypothetical prote
35	21	100.0	118	2	T38861	small nuclear ribo
36	21	100.0	119	2	B32578	T-cell receptor be
37	21	100.0	120	2	S57887	T cell receptor CK
38	21	100.0	121	2	S09793	hypothetical prote
39	21	100.0	121	2	D69031	hypothetical prote
40	21	100.0	122	2	S57888	T cell receptor Bb
41	21	100.0	123	2	G64219	ribosomal protein
42	21	100.0	123	2	B64418	conserved hypotnet
43	21	100.0	129	2	T05491	hypothetical prote
44	21	100.0	129	2	S14984	glycine-rich prote
45	21	100.0	130	2	B72531	hypothetical prote

ALIGNMENTS

RESULT 1

A44773  
pollen allergen I - Japanese cedar (fragment)  
C;Species: Cryptomeria japonica (Japanese cedar)  
C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 11-Jan-2000  
C;Accession: A44773  
R;Tanai, M.; Ando, S.; Usui, M.; Kurimoto, M.; Sakaguchi, M.; Inouye, S.; Matukasi, T.  
FEBS Lett. 239, 329-332, 1988  
A;Title: N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (C)  
A;Reference number: A44773; MUID:89031257; PMID:3181436  
A;Accession: A44773  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <TAN>  
C;Superfamily: pectate lyase LAT59  
C;Keywords: pollen

Query Match 100.0%; Score 21; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 9 RGDS 12

RESULT 2

S15982  
arietin - puff adder (fragment)  
C;Species: Bitis arietans (puff adder)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 29-May-1998  
C;Accession: S15982; S15983  
R;Huang, T.F.; Wang, W.J.; Teng, C.M.; Liu, C.S.; Ouyang, C.  
Biochim. Biophys. Acta 1074, 136-143, 1991  
A;Title: Purification and characterization of an antipeptide, arietin, from Bit  
A;Reference number: S15982; MUID:91255281; PMID:2043663  
A;Accession: S15982  
A;Molecule type: protein  
A;Residues: 1-30 <HUA>  
C;Superfamily: disintegrin homology  
C;Keywords: venom

Query Match 100.0%; Score 21; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 24 RGDS 27

RESULT 3

D82489  
 hypothetical protein VCA0187 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C:Species: Vibrio cholerae  
 C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: D82489  
 R:Heidelberger, J.F.; Eissen, J.A.; Nelson, M.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardac, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: D82489  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-43 <HEI>  
 A:Cross-references: GB:AE004359; GB:AE003853; NID:99657575; PIDN:AAF96100.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VCA0187  
 A:Map position: 2

Query Match 100.0%; Score 21; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 DB 10 RGDS 13

## RESULT 4

S69113  
 collagen alpha 3(IV) chain, splice form GP-III/IV/V - human  
 C:Species: Homo sapiens (man)  
 C>Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 12-Nov-1999  
 C:Accession: S69113  
 R:Penades, J.R.; Bernal, D.; Revert, F.; Johansson, C.; Fresquet, V.J.; Cervera, J.; Wiedt, J. Biochem. 229, 754-760, 1995  
 A:Title: Characterization and expression of multiple alternatively spliced transcripts of alpha 3(IV) chain and one of its alternative forms.  
 A:Reference number: S69111; MUID:95278230; PMID:7758473  
 A:Accession: S69113  
 A:Molecule type: mRNA  
 A:Residues: 1-52 <PEN>  
 C:Comment: For the complete sequence of the long splice form, see PIR:CGHUB.  
 C:Superfamily: collagen alpha 1(IV) chain

Query Match 100.0%; Score 21; DB 2; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 DB 15 RGDS 18

## RESULT 5

S03486  
 T-cell receptor beta chain V-D-J region (clone HBP04) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-May-1997  
 C:Accession: S03486  
 R:Kimura, N.; Toyonaga, B.; Yoshikai, Y.; Triebel, F.; Debre, P.; Minder, M.D.; Mak, T.W. J. Exp. Med. 164, 739-750, 1986  
 A:Title: Sequences and diversity of human T cell receptor beta chain variable region genes  
 A:Reference number: S03485; MUID:86306525; PMID:3755748  
 A:Accession: S03486  
 A:Molecule type: mRNA  
 A:Residues: 1-53 <KIN>  
 A:Cross-references: EMBL:X04922  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: T-cell receptor

Query Match 100.0%; Score 21; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 DB 15 RGDS 18

## RESULT 6

S18571  
 otcZ protein - Streptomyces rimosus (fragment)  
 C:Species: Streptomyces rimosus  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
 C:Accession: S18571  
 R:Doyle, D.; McDowell, K.J.; Butler, M.J.; Hunter, I.S. Mol. Microbiol. 5, 2923-2933, 1991  
 A:Title: Characterization of an oxytetracycline-resistance gene, otcA, of Streptomyces  
 A:Reference number: S18571; MUID:92236410; PMID:1809836  
 A:Accession: S18571  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-62 <DOY>  
 A:Cross-references: EMBL:X53401  
 C:Genetics:  
 A:Gene: otcZ

Query Match 100.0%; Score 21; DB 2; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 DB 20 RGDS 23

## RESULT 7

AF2139  
 hypothetical protein asr2669 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AF2139  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Katayabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, N. DNA Res. 8, 203-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AF2139  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-88 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BAW74368.1; PID:gl7131762; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: asr2669

Query Match 100.0%; Score 21; DB 2; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 DB 30 RGDS 33

## RESULT 8

C90091  
 small nucleolar ribonucleoprotein G [imported] - Giardia theta nucleomorph  
 C:Species: nucleomorph Giardia theta  
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 C:Accession: C90091

R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil Nature 410, 1037-1096, 2001  
 A>Title: The highly reduced genome of an enslaved algal nucleus.  
 A;Reference number: A99082; MUID:11323671; PMID:11323671

A;Accession: C90091  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-74 <DOU>  
 A;Cross-references: GB:AF165818; NID:gl3794485; PIDN:AAK39863.1; GSPDB:GN00150  
 C;Genetics:  
 A;Gene: snrpg  
 A;Map position: 1  
 A;Genome: nucleomorph  
 C;Keywords: nucleomorph

Query Match 100.0%; Score 21; DB 2; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 62 RGDS 65

RESULT 9  
 D82844  
 Carbon storage regulator XF0125 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
 C;Accession: D82844

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
 Nature 406, 152-157, 2000  
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: D82844  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-76 <SIM>

A;Cross-references: GB:AE003866; GB:AE003845; NID:99104906; PIDN:AAF92938.1; GSPDB:GN001  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B  
 B;Bioness, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
 as-Neto, E.; Decena, C.; El-Porcy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigh  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0125

C;Superfamily: glycogen biosynthesis inhibitor  
 Query Match 100.0%; Score 21; DB 2; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 69 RGDS 72

RESULT 10  
 G72749  
 probable small nuclear ribonucleoprotein APES022 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C;Accession: G72749

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takak  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Taraka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: G72749  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-77 <KAW>

A;Cross-references: DDBJ:AP000059; NID:95103911; PIDN:BAA79431.1; PID:dl043277; PID:9510  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APES022

Query Match 100.0%; Score 21; DB 2; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 66 RGDS 69

RESULT 11  
 S17389  
 T-cell receptor beta chain V region (clone IGRb12) - human (fragment)  
 C;Species: Homo sapiens (man)  
 C;Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
 C;Accession: S17389

R;Ferradini, L.; Roman-Roman, S.; Azocar, J.; Michalaki, H.; Triebel, F.; Hercend, T.  
 Eur. J. Immunol. 21, 935-942, 1991  
 A;Title: Studies on the human T cell receptor alpha/beta variable region genes. II. Ider  
 A;Reference number: S17378; MUID:91209402; PMID:1826889

A;Accession: S17389  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-79 <FER>

A;Cross-references: EMBL:X58807; NID:933545; PIDN:CAA41613.1; PID:9312204  
 C;Superfamily: immunoglobulin V region; immunoglobulin homo-cg  
 C;Keywords: T-cell receptor  
 F;1-78/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 21; DB 2; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 68 RGDS 71

RESULT 12  
 A49736  
 collagen alpha 3(IV) chain, short splice form - human (fragment)  
 N;Contains: collagen alpha 3 (IV) chain, splice form GP-III  
 C;Species: Homo sapiens (man)

C;Date: 03-May-1994 #sequence\_revision 12-Nov-1999 #text\_change 12-Nov-1999  
 C;Accession: A49736; C49736; B45971; S69112  
 R;Peng, L.; Xia, Y.; Wilson, C.B.  
 J. Biol. Chem. 269, 2342-2348, 1994

A;Title: Alternative splicing of the NCI domain of the human alpha3(IV) collagen gene. E  
 A;Reference number: A49736; MUID:94124597; PMID:8294492

A;Accession: A49736  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 71-81 <FEN1>

A;Accession: C49736  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 22-81 <FEN2>

A;Cross-references: GB:U02520; NID:9408895; PIDN:AAA18943.1; PID:9408896  
 A;Note: this is the conceptual translation of the nucleic acid submitted to GenBank

R;Bernal, D.; Quinones, S.; Saus, J.  
 J. Biol. Chem. 268, 12090-12094, 1993  
 A;Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
 A;Reference number: A45971; MUID:93280184; PMID:8505352  
 A;Accession: B45971  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 71-81 <BER>  
 A;Cross-references: PIDN:AB27014.1; 2D:G385563  
 A;Note: sequence extracted from NCBI backbone (NCBIP:133955); sequence incorrectly identified  
 R;Penades, J.R.; Bernal, D.; Revert, F.; Johansson, C.; Fresquet, V.J.; Cervera, J.; Wise  
 Eur. J. Biochem. 229, 754-760, 1995  
 A;Title: Characterization and expression of multiple alternatively spliced transcripts of  
 uroantigen and one of its alternative forms.  
 A;Reference number: S69111; MUID:95278230; PMID:7758473  
 A;Accession: S69112  
 A;Molecule type: mRNA  
 A;Residues: 1-45,71-81 <PEN>  
 A;Comment: For the complete sequence of the long splice form, see PIR:CSHU3B.  
 C;Genetics:  
 A;Gene: GDB:COL4A3  
 A;Cross-references: GDB:128351; OMIM:120070  
 A;Map position: 2q36-2q37  
 C;Superfamily: collagen alpha 1(IV) chain  
 C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracellular  
 F;1-81/Product: collagen alpha 3(IV) chain, short splice form (fragment) #status predict  
 F;1-45,71-81/Product: collagen alpha 3 (IV) chain, splice form GP-III (fragment) #status  
 F;22-81/Domain: carboxyl-terminal nonhelical, NCI <NCI>

Query Match 100.0%; Score 21; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 15 RGDS 18

## RESULT 13

AC1886  
 hypothetical protein asr0636 [imported] - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of *Arabaena* sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: AC1886  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watarabe, A.; Iriguchi,  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 9, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AC1886  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-84 <KUR>  
 A;Cross-references: GB:BA000013; PIDN:BA072594.1; PID:gl7129982; GSPDB:GN00179  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: asr0636

Query Match 100.0%; Score 21; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 16 RGDS 19

## RESULT 14

JQ1683  
 UL49A protein - human herpesvirus 2  
 C;Species: human herpesvirus 2  
 C;Date: 10-Jun-1993 #sequence\_revision 10-Jun-1993 #text\_change 21-Jul-2000

C;Accession: JQ1683  
 R;Barnett, B.C.; Dolan, A.; Telford, E.A.R.; Davison, A.J.; McGeech, D.J.  
 J. Gen. Virol. 73, 2167-2171, 1992  
 A;Title: A novel herpes simplex virus gene (UL49A) encodes a putative membrane protein  
 A;Reference number: JQ1683; MUID:92356101; PMID:1322965  
 A;Accession: JQ1683  
 A;Molecule type: DNA  
 A;Residues: 1-87 <BAR>  
 A;Cross-references: GB:286099; NID:G6572414; PIDN:CAB06736.1; PID:gl869873  
 A;Experimental source: strain HG52

Query Match 100.0%; Score 21; DB 2; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 43 RGDS 46

## RESULT 15

Tl4944  
 hypothetical protein Y1028 - *Yersinia pestis* plasmid pMT1  
 C;Species: *Yersinia pestis*  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 22-Oct-1999  
 C;Accession: Tl4944; Tl4671  
 R;Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.  
 Infect. Immun. 66, 5731-5742, 1998  
 A;Title: Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIMS plasmid  
 A;Reference number: Z18268; MUID:99043898; PMID:9826348  
 A;Accession: Tl4944  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-88 <LIN>  
 A;Cross-references: EMBL:AF074611; NID:G3883003; PID:G3883029; PIDN:AAC52689.1  
 R;Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrano  
 submitted to the EMBL Data Library, March 1998  
 A;Description: Structural organization of virulence determinants in three *Yersinia pestis*  
 A;Reference number: Z18168  
 A;Accession: Tl4671  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-88 <HUP>  
 A;Cross-references: EMBL:AF053947; NID:G2996286; PID:G2996306; PIDN:AAC13186.1  
 C;Genetics:  
 A;Gene: Y1028  
 A;Genome: plasmid pMT1

Query Match 100.0%; Score 21; DB 2; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 63 RGDS 66

Search completed: April 16, 2004, 07:29:26  
 Job time : 9.4 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:55:57 ; Search time 43.7 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588b-23

Perfect score: 21

Sequence: 1 RGDs 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jard4:\*

1: Geneseq1990s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2000s:\*

5: Geneseq2000s:\*

6: Geneseq2000s:\*

7: Geneseq2000s:\*

8: Geneseq2000s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	21	100.0	4	1 AAP61325	Aap61325 Sequence
2	21	100.0	4	2 AAR04608	Aar04608 Antiviral
3	21	100.0	4	2 AAR26802	Aar26802 Triglycer
4	21	100.0	4	2 AAR25312	Aar25312 Cell cont
5	21	100.0	4	2 AAR26810	Aar26810 Polyethyl
6	21	100.0	4	2 AAR26806	Aar26806 Propan-am
7	21	100.0	4	2 AAR26398	Aar26398 Sequence
8	21	100.0	4	2 AAR36713	Aar36713 Adhesion
9	21	100.0	4	2 AAR37833	Aar37833 Cell adhe
10	21	100.0	4	2 AAR44043	Aar44043 RGD pepti
11	21	100.0	4	2 AAR37134	Aar37134 RGD pepti
12	21	100.0	4	2 AAR37133	Aar37133 RGD pepti
13	21	100.0	4	2 AAR37137	Aar37137 RGD pepti
14	21	100.0	4	2 AAR37138	Aar37138 RGD pepti
15	21	100.0	4	2 AAR37132	Aar37132 RGD pepti
16	21	100.0	4	2 AAR37135	Aar37135 RGD pepti
17	21	100.0	4	2 AAR37130	Aar37130 RGD pepti
18	21	100.0	4	2 AAR37136	Aar37136 RGD pepti
19	21	100.0	4	2 AAR35464	Aar35464 Propene-a
20	21	100.0	4	2 AAR44667	Aar44667 Platelet
21	21	100.0	4	2 AAR67015	Aar67015 RGD pepti
22	21	100.0	4	2 AAR57828	Aar57828 RGD contg
23	21	100.0	4	2 AAR53729	Aar53729 Tetrapept
24	21	100.0	4	2 AAR85667	Aar85667 Anti-alle
25	21	100.0	4	2 AAR62946	Aar62946 RGD contg

26	21	100.0	4	2 AAR70473	Aar70473 Cancer me
27	21	100.0	4	2 AAR29563	Aar29563 RGD pepti
28	21	100.0	4	2 ADD94859	Add94859 Platelet
29	21	100.0	4	2 AAW11092	Aaw11092 Platelet-
30	21	100.0	4	2 AAR98812	Aar98812 Cell adhe
31	21	100.0	4	2 AAW25174	Aaw25174 RGD-pepti
32	21	100.0	4	2 AAW45489	Aaw45489 Targeting
33	21	100.0	4	2 AAW45339	Aaw45339 Peptide f
34	21	100.0	4	2 AAW35541	Aaw35541 Peptide S
35	21	100.0	4	2 AAW31144	Aaw31144 Platelet-
36	21	100.0	4	2 AAW62917	Aaw62917 Part of t
37	21	100.0	4	2 AAW66842	Aaw66842 Peptide u
38	21	100.0	4	2 AAW63134	Aaw63134 Peptide s
39	21	100.0	4	2 AAW26953	Aaw26953 RGD-conta
40	21	100.0	4	2 AAY16022	Aay16022 Snake ven
41	21	100.0	4	2 AAW26951	Aaw26951 RGD-conta
42	21	100.0	4	2 AAW84460	Aaw84460 RGD pepti
43	21	100.0	4	2 AAY50313	Aay50313 Neutroph
44	21	100.0	4	2 AAY07267	Aay07267 RGDs pept
45	21	100.0	4	2 AAY22413	Aay22413 Fibronect

## ALIGNMENTS

RESULT 1  
AAP61325  
ID AAP61325 standard; peptide; 4 AA.  
XX AAP61325;  
AC AC

DT 03-OCT-2002 (revised)  
DT 22-JUN-1991 (first entry)  
XX  
XX

DE Sequence of peptide with the same cell-attachment-promoting activity as  
DE fibronectin.

KW Cell attachment; phagocytosis; injury therapy.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "attached to H or at least one AA"

FT Misc-difference 4

FT Misc-difference /note= "attached to OH or at least one AA"

XX US4614517-A.

PD 30-SEP-1986.

PF 17-JUN-1985; 85US-00745086.

PR 04-AUG-1982; 82US-00405239.

PR 08-OCT-1982; 82US-00433457.

PR 28-JUL-1983; 83US-00518036.

PR 22-NOV-1983; 83US-00554821.

XX WPI; 1986-278506/42.

XX Claim 1; Col 8; 8pp; English.

CC The patentors claim a compsn. which comprises AAP61325 is useful in  
CC surgery and therapeutic reconstruction and treatment of injuries.  
CC AAP61325 is the tetrapeptide sequence of fibronectin which promotes  
CC attachment of suspended cells to tissue culture substrate as well as to  
CC collagen, and is common to other proteins which interact with cells.  
CC (Updated on 03-OCT-2002 to add missing CS field.)

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

RESULT 2  
AAR04608  
ID AAR04608 standard; protein; 4 AA.  
XX AC AAR04608;  
XX DT 25-MAR-2003 (revised)  
DT 05-SEP-1990 (first entry)  
XX DE Antiviral agent.  
XX KW Antiviral; M2; poliovirus; polio; hepatitis.  
XX OS Synthetic.  
XX PV JP02078631-A.  
XX PD 19-MAR-1990.  
XX PF 14-SEP-1988; 88JP-00228843.  
XX ER 14-SEP-1988; 88JP-00228843.  
XX PA (NIHA ) NIPPON MINING CO.  
XX WPI; 1990-123060/17.  
XX Antiviral agent contg. tri:peptide (unit) - of basic aminoacid, then  
PT alanine, glycine or sarcosine, and acidic aminoacid, effective against  
PT virus with protein-terminated DNA or RNA.  
PS Disclosure; Page ?; 4pp; Japanese.  
CC Peptide is effective against inhibiting propagation of DNA or RNA bonded,  
CC protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on  
CC 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

RESULT 3  
AAR26802  
ID AAR26802 standard; peptide; 4 AA.  
XX AC AAR26802;  
XX DT 11-FEB-1993 (first entry)  
XX DE Triglyceride-substituted peptide #1.  
XX KW Cell movement inhibitor; cell culture; adhesion membrane.  
XX OS Synthetic.  
XX PV Key Location/Qualifiers  
FT Modified-site 1 /note= "Arg is N-acetylated by the group; sn-CH2(OR1)-

Best Local Similarity 100.0%; Score 21; DB 2; Length 4;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

RESULT 4  
AAR25312  
ID AAR25312 standard; peptide; 4 AA.  
XX AC AAR25312;  
XX DT 17-MAR-1993 (first entry)  
XX DE Cell contact inhibitor generic peptide #1.  
XX KW Cyclic peptide; cell contact inhibitor; hydrolytic enzyme.  
XX OS Synthetic.  
XX PV Key Location/Qualifiers  
FT Modified-site 2 /label= MeGly  
XX PN JP04264097-A.  
XX PD 18-SEP-1992.  
XX PF 16-FEB-1991; 91JP-00044386.  
XX PR 16-FEB-1991; 91JP-00044386.  
XX PA (ASAG ) ASahi GLASS CO LTD.  
XX WPI; 1992-361922/44.  
XX Peptide derivs. as contact inhibitor for animal cells - comprise  
PT synthesised cyclic peptide and have portion of aminoacid sequence of  
PT arginine-N-methyl:glycine-aspartic acid.  
XX PS Disclosure; Page 3; 6pp; Japanese.  
XX CC The sequences given in AAR25311-19 are cyclic peptides which act as  
CC contact inhibitors of animal cells. They are resistant to decomposition

CH(OR1)-CH2-O-CO(CH2)2-CO-Arg where R1= C16H33, n-C15H31  
or 3,7,11,15-tetramethyl-hexadecyl"

JP04208296-A.  
29-JUL-1992.  
30-NOV-1990; 90JP-00334794.  
30-NOV-1990; 90JP-00334794.  
(FUJIF ) FUJIFILM PHOTO FILM CO LTD.  
WPI; 1992-303598/37.  
Peptide deriv. - useful as cell movement inhibitor, cell adhesion  
membrane or as culture substrate.  
Disclosure; Page 3; 8pp; Japanese.  
The sequences given in AAR26802-4 are examples of a highly generic  
claimed peptide. These peptides are useful as cell movement inhibitors,  
cell adhesion membranes and cell culture substrates  
Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

RESULT 4  
AAR25312  
ID AAR25312 standard; peptide; 4 AA.  
XX AC AAR25312;  
XX DT 17-MAR-1993 (first entry)  
XX DE Cell contact inhibitor generic peptide #1.  
XX KW Cyclic peptide; cell contact inhibitor; hydrolytic enzyme.  
XX OS Synthetic.  
XX PV Key Location/Qualifiers  
FT Modified-site 2 /label= MeGly  
XX PN JP04264097-A.  
XX PD 18-SEP-1992.  
XX PF 16-FEB-1991; 91JP-00044386.  
XX PR 16-FEB-1991; 91JP-00044386.  
XX PA (ASAG ) ASahi GLASS CO LTD.  
XX WPI; 1992-361922/44.  
XX Peptide derivs. as contact inhibitor for animal cells - comprise  
PT synthesised cyclic peptide and have portion of aminoacid sequence of  
PT arginine-N-methyl:glycine-aspartic acid.  
XX PS Disclosure; Page 3; 6pp; Japanese.  
XX CC The sequences given in AAR25311-19 are cyclic peptides which act as  
CC contact inhibitors of animal cells. They are resistant to decomposition

CC by hydrolytic enzymes and can be maintained at high levels of activity  
 CC for a long period in vivo. The peptides are cyclic and may have 1-16  
 CC pref. 1-4 amino acids

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 1 RGDS 4

## RESULT 5

AAR26810  
 ID AAR26810 standard; peptide; 4 AA.

XX AC AAR26810;

XX DT 20-MAY-1998 (first entry)

XX DE Polyethylene glycol derivative #1.

XX KW triazine ring; methoxy-polyethyleneoxy group; fibronectin; vitronectin;  
 XX platelet adhesion; metastasis; neuropathy.

XX OS Synthetic.

XX PN JP04217693-A.

XX PD 07-JUG-1992.

XX PF 30-NOV-1990; 90JP-00333717.

XX PR 23-OCT-1990; 90JP-00285172.

XX PA (FUJIF ) FUJIF PHOTO FILM CO LTD.

XX DR WPI; 1992-312284/38.

XX PT Polyethylene glycol derivs. contg. peptide(s) - inhibit cellular adhesion  
 XX for fibronectin or vitronectin and are used to inhibit agglutination or  
 XX adhesion of platelets.

XX PS Disclosure; Page 3; 9pp; Japanese.

XX CC The sequences given in AAR26810-14 are examples of a peptide chain which  
 CC is attached once or twice to a triazine ring which is also substituted  
 CC twice or once, respectively, with a methoxy-polyethyleneoxy group. These  
 CC peptides can be used to inhibit cellular adhesion to fibronectin or  
 CC vitronectin and they are useful as inhibitors for agglutination or  
 CC adhesion of platelets. They can also be useful as inhibitors for  
 CC metastasis of cancers, inhibitors of agglutination of platelets caused by  
 CC tumour cells in the blood capillaries, and drugs acting on neuropathy

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 1 RGDS 4

## RESULT 6

AAR26806

ID AAR26806 standard; peptide; 4 AA.

XX AC AAR26806;

XX DT 12-FEB-1993 (first entry)  
 XX DE Propen-amido peptide copolymer.  
 XX KW Tumour metastasis inhibitor; platelet aggregation; animal cell; adhesion;  
 XX wound healing; cell culture media.  
 XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX Modified-site 1 /note= "CH2=CH-CO-NH-(CH2)3-CO-Arg"

XX PN JP04213311-A.

XX PD 04-AUG-1992.

XX PF 29-MAR-1991; 91JP-00066159.

XX PR 27-NOV-1990; 90JP-00324610.

XX PA (FUJIF ) FUJIF PHOTO FILM CO LTD.

XX DR WPI; 1992-305482/37.

XX CC New copolymers of propen-amido peptide(s) - are tumour metastasis,  
 XX platelet aggregation and animal cell adhesion inhibitors also useful as  
 XX wound healing agents and cell culture media.

XX PS Example; Page 6; 14pp; Japanese.

XX CC The sequence is that of a propen-amido peptide copolymer, it and its  
 XX salts are water soluble and is useful as a tumour metastasis inhibitor, a  
 XX platelet aggregation inhibitor, an animal cell adhesion inhibitor, a  
 XX wound healing agent and cell culture media. See also AAR26805-R26808

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
 Db 1 RGDS 4

## RESULT 7

AAR26398

ID AAR26398 standard; peptide; 4 AA.

XX AC AAR26398;

XX DT 25-MAR-2003 (revised)

XX DT 25-JAN-1993 (first entry)

XX DE Sequence of platelet binding peptide.

XX KW Technetium-99m labelled polypeptide imaging agent; radiolabeled imaging;  
 XX radiodiagnostic agent.

XX OS Synthetic.

XX PN WO9213572-A1.

XX PD 20-AUG-1992.

XX PF 07-FEB-1992; 92WO-US000757.

XX PR 08-FEB-1991; 91US-00653012.

XX PA (DIAT-) DIATECH INC.

XX  
PI Dean RT;  
XX  
DR WPI; 1992-233767/36.  
XX  
PT New technetium-99m labelled polypeptide imaging agents - for imaging of  
PT clots, tumours, infection sites, atherosclerotic and amyloid plaques or  
PT bone, and for visualising organs.  
XX  
PS Claim 6; Page 13; 19pp; English.  
XX  
CC The binding peptide is covalently linked to a 'Op(aa)Cp' technetium  
CC binding group wherein Cp is a protected cysteine and (aa) is an amino  
CC acid. The technetium-99m complexes are used to image target sites within  
CC a mammalian body. (Updated on 25-MAR-2003 to correct EN field.)  
XX  
XX Sequence 4 AA;  
SQ  
Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;  
QY 1 RGDS 4  
DB 1 RGDS 4  
|||||  
XX  
RESULT 8  
AAR36713  
ID AAR36713 standard; peptide; 4 AA.  
XX  
AC AAR36713;  
XX  
XX 25-MAR-2003 (revised)  
DT 26-AUG-1993 (first entry)  
XX  
XX Adhesion formation prevention RGD-contg. peptide.  
DE  
XX Tissue repair; peritoneum; surgery; post-surgically; inhibition;  
KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;  
KW CNS; use.  
XX  
XX Synthetic.  
XX  
XX WO9308818-A1.  
XX  
XX 13-MAY-1993.  
XX  
XX 06-NOV-1992; 92WO-US009494.  
XX  
XX 07-NOV-1991; 91US-00789231.  
XX  
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
XX  
XX Dizerega GS, Rodgers KE;  
XX  
XX WPI; 1993-167381/20.  
XX  
XX Prevention of adhesion formation, partic. post-surgically - comprises  
PT administering a RGD-contg. peptide for a time sufficient to permit tissue  
PT repair.  
XX  
XX Example; Page 18; 22pp; English.  
XX  
XX The sequence is that of an RGD-contg. peptide which is used in a method  
CC for prevention of adhesion formation for a time sufficient to permit  
CC tissue repair. The method is used for minimising or preventing adhesion  
CC formation, partic. in the peritoneum following surgery, but also for e.g.  
CC cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In  
CC addn., the peptide inhibits platelet aggregation and does not induce  
CC inflammation or trauma at the site of administration. (Updated on 25-MAR-  
CC 2003 to correct PN field.)  
XX

SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDS 4  
DB 1 RGDS 4  
|||||  
XX  
RESULT 9  
AAR37833  
ID AAR37833 standard; protein; 4 AA.  
XX  
AC AAR37833;  
XX  
XX 12-OCT-1993 (first entry)  
DT  
XX Cell adhesion motif encoded by insert 74RGD4.  
DE  
XX Fibronectin; cell-cell adhesion; Arg-Gly-Asp motif; human lysozyme;  
KW bacteriolysis.  
XX  
XX Synthetic.  
XX  
XX JP05115284-A.  
XX  
XX 14-MAY-1993.  
PD  
XX 29-OCT-1991; 91JP-00282866.  
PF  
XX 29-OCT-1991; 91JP-00282866.  
PR  
XX (FUJI-) FUCITA GAKUEN GH.  
XX  
XX (TANP-) TANPAKU KOGAKU KENKYUSHO KK.  
PA  
XX WPI; 1993-190693/24.  
DR  
XX N-PSDB; AAQ42607.  
XX  
XX New mutant human lysozyme, - is incorporated into chimeric protein for  
PT use in analysing high order structure of aminoacid sequence contg. cell  
PT adhesive function site.  
XX  
XX Example 1; Page 6; 10pp; Japanese.  
XX  
XX The double-stranded DNA 74RGD4 is one of 5 preferred sequences, each  
CC coding for a cell-adhesion motif, which can be inserted between the  
CC codons for Val 74 and Asn 75 of human lysozyme. The resulting mutant  
CC lysozyme polypeptide has cell adhesion properties in addition to its  
CC bacteriolytic ability. See also AAQ42608-Q42611  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGDS 4  
DB 1 RGDS 4  
|||||  
XX  
RESULT 10  
AAR44043  
ID AAR44043 standard; peptide; 4 AA.  
XX  
AC AAR44043;  
XX  
XX 02-JUN-1994 (first entry)  
DT  
XX RGD peptide derivative #1.  
XX  
DE  
XX



CC derivs. contain the cell adhesion core sequence Arg- Gly-Asp. They are  
 CC useful as inhibitors of cancer metastasis

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 1 RGDS 4

## RESULT 13

AAR37137  
 ID AAR37137 standard; peptide; 4 AA.

XX AAR37137;

DT 19-APR-1994 (first entry)

XX RGD peptide deriv. #8.

DE Cell adhesion core sequence; inhibitor; cancer; metastasis; ss.

XX Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 4 /label= Suc-D-Arg

FT Modified-site 4 /label= Ser-OMe

EN JP05222092-A.

XX 31-AUG-1993.

XX 16-JUL-1992; 92JP-00189446.

XX 18-DEC-1991; 91JP-00335213.

XX (FUJF ) FUJII PHOTO FILM CO LTD.

XX WPI; 1993-309192/39.

PT Peptide derivs. used as cancer metastasis inhibitors - contain the cell  
 PT adhesion core sequence Arg-Gly-Asp followed by Ser.

XX Disclosure; Page 3; 10pp; Japanese.

CC The sequences given in AAR37130-48 are peptide derivatives which  
 CC correspond to the generic formula; R1-X-Arg-Gly-Asp-Ser-R2; R1 = H or an  
 CC optionally substituted acyl, R2 = -OR3 or -NR4R5, R3, R4, R5 = H or 1-3C  
 CC alkyl, and X = a bond or an amino acid or peptide residue. These peptide  
 CC derivs. contain the cell adhesion core sequence Arg- Gly-Asp. They are  
 CC useful as inhibitors of cancer metastasis

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 1 RGDS 4

## RESULT 14

AAR37138

ID AAR37138 standard; peptide; 4 AA.

XX

AC AAR37138;  
 XX 19-APR-1994 (first entry)  
 DT RGD peptide deriv. #9.  
 DE Cell adhesion core sequence; inhibitor; cancer; metastasis; ss.  
 KW Cell adhesion core sequence; inhibitor; cancer; metastasis; ss.  
 XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= Adipoyl-Arg

XX JP05222092-A.

XX 31-AUG-1993.

XX 16-JUL-1992; 92JP-00189446.

XX 18-DEC-1991; 91JP-00335213.

XX (FUJF ) FUJII PHOTO FILM CO LTD.

XX WPI; 1993-309192/39.

XX Peptide derivs. used as cancer metastasis inhibitors - contain the cell  
 XX adhesion core sequence Arg-Gly-Asp followed by Ser.

XX Disclosure; Page 3; 10pp; Japanese.

CC The sequences given in AAR37130-48 are peptide derivatives which  
 CC correspond to the generic formula; R1-X-Arg-Gly-Asp-Ser-R2; R1 = H or an  
 CC optionally substituted acyl, R2 = -OR3 or -NR4R5, R3, R4, R5 = H or 1-3C  
 CC alkyl, and X = a bond or an amino acid or peptide residue. These peptide  
 CC derivs. contain the cell adhesion core sequence Arg- Gly-Asp. They are  
 CC useful as inhibitors of cancer metastasis

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 ||||  
 Db 1 RGDS 4

## RESULT 15

AAR37132

ID AAR37132 standard; peptide; 4 AA.

XX AAR37132;

DT 19-APR-1994 (first entry)

XX RGD peptide deriv. #3.

DE Cell adhesion core sequence; inhibitor; cancer; metastasis; ss.

XX Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= Suc-D-Arg

XX JP05222092-A.

XX 31-AUG-1993.

XX 16-JUL-1992; 92JP-00189446.

PR 18-DEC-1991; 9LJP-00335213.  
 XX  
 PA (FUJIF ) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1993-309192/39.  
 XX  
 PT Peptide derivs. used as cancer metastasis inhibitors - contain the cell  
 PT adhesion core sequence Arg-Gly-Asp followed by Ser.  
 XX  
 PS Disclosure; Page 3; 10pp; Japanese.  
 XX  
 CC The sequences given in AAR37130-48 are peptide derivatives which  
 CC correspond to the generic formula: R1-X-Arg-Gly-Asp-Ser-R2; R1 = H or an  
 CC optionally substituted acyl, R2 = -OR3 or -NR4R5, R3, R4, R5 = H or 1-3C  
 CC alkyl, and X = a bond or an amino acid or peptide residue. These peptide  
 CC derivs. contain the cell adhesion core sequence Arg- Gly-Asp. They are  
 CC useful as inhibitors of cancer metastasis  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDS 4  
 ||||  
 Db 1 RGDS 4

Search completed: April 16, 2004, 07:20:57  
 Job time : 44.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 12.9 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-23  
Perfect score: 21  
Sequence: 1 RGDS 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2.6/prodata/2/iaa/5A COMB pep.\*  
2: /cgn2.6/prodata/2/iaa/5B COMB pep.\*  
3: /cgn2.6/prodata/2/iaa/6A COMB pep.\*  
4: /cgn2.6/prodata/2/iaa/6B COMB pep.\*  
5: /cgn2.6/prodata/2/iaa/PCUS COMB pep.\*  
6: /cgn2.6/prodata/2/iaa/backfiles1 pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	21	100.0	4	1 US-07-780-790A-7	Sequence 7, Appl
2	21	100.0	4	1 US-07-683-957B-7	Sequence 7, Appl
3	21	100.0	4	1 US-07-834-848-16	Sequence 16, Appl
4	21	100.0	4	1 US-08-127-351-54	Sequence 54, Appl
5	21	100.0	4	1 US-08-109-106-1	Sequence 1, Appl
6	21	100.0	4	1 US-08-251-027-8	Sequence 8, Appl
7	21	100.0	4	1 US-08-480-367B-54	Sequence 54, Appl
8	21	100.0	4	1 US-08-487-221A-54	Sequence 54, Appl
9	21	100.0	4	1 US-08-480-370-54	Sequence 54, Appl
10	21	100.0	4	1 US-08-389-035-1	Sequence 1, Appl
11	21	100.0	4	1 US-08-178-482-11	Sequence 11, Appl
12	21	100.0	4	1 US-08-299-636-35	Sequence 35, Appl
13	21	100.0	4	1 US-08-279-155-34	Sequence 34, Appl
14	21	100.0	4	1 US-08-142-449B-13	Sequence 13, Appl
15	21	100.0	4	1 US-08-464-456-33	Sequence 33, Appl
16	21	100.0	4	1 US-08-128-225-2	Sequence 2, Appl
17	21	100.0	4	1 US-08-405-200-1	Sequence 1, Appl
18	21	100.0	4	1 US-08-596-116A-67	Sequence 67, Appl
19	21	100.0	4	1 US-08-596-116A-70	Sequence 70, Appl
20	21	100.0	4	1 US-08-338-282-12	Sequence 12, Appl
21	21	100.0	4	1 US-08-658-130-18	Sequence 18, Appl
22	21	100.0	4	1 US-08-462-661A-108	Sequence 108, Appl
23	21	100.0	4	1 US-08-421-702A-28	Sequence 28, Appl
24	21	100.0	4	1 US-08-343-264-1	Sequence 1, Appl
25	21	100.0	4	1 US-08-487-568-55	Sequence 55, Appl
26	21	100.0	4	1 US-08-421-696A-28	Sequence 28, Appl
27	21	100.0	4	1 US-08-703-988A-34	Sequence 34, Appl

28	21	100.0	4	1 US-08-463-052-33	Sequence 33, Appl
29	21	100.0	4	1 US-08-421-697A-28	Sequence 28, Appl
30	21	100.0	4	1 US-08-421-698A-28	Sequence 28, Appl
31	21	100.0	4	1 US-08-668-871-13	Sequence 13, Appl
32	21	100.0	4	2 US-08-480-551-33	Sequence 33, Appl
33	21	100.0	4	2 US-08-669-683-13	Sequence 13, Appl
34	21	100.0	4	2 US-08-836-854-1	Sequence 1, Appl
35	21	100.0	4	2 US-08-480-133A-14	Sequence 14, Appl
36	21	100.0	4	2 US-08-421-695A-28	Sequence 28, Appl
37	21	100.0	4	2 US-08-807-464-1	Sequence 1, Appl
38	21	100.0	4	2 US-08-934-222-6	Sequence 6, Appl
39	21	100.0	4	2 US-08-747-137-25	Sequence 25, Appl
40	21	100.0	4	2 US-08-747-137-180	Sequence 180, Appl
41	21	100.0	4	2 US-08-933-402-6	Sequence 6, Appl
42	21	100.0	4	2 US-09-207-621-6	Sequence 6, Appl
43	21	100.0	4	2 US-08-532-818-6	Sequence 6, Appl
44	21	100.0	4	2 US-08-612-842-34	Sequence 34, Appl
45	21	100.0	4	2 US-08-574-699A-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-07-780-790A-7  
; Sequence 7, Application US/07780790A  
; Patent No. 5298488  
; GENERAL INFORMATION:  
; APPLICANT: KOJIMA, Masayoshi  
; APPLICANT: KOMAZAWA, Hiroyuki  
; TITLE OF INVENTION: CM-CHITIN DERIVATIVES AND USE THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak, & Sess  
; STREET: 210C Pennsylvania Avenue  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: USA  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07780,790A  
; FILING DATE: 19911023  
; CLASSIFICATION: 530  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-780-790A-7

Query Xmatch 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDS 4  
DB 1 RGDS 4

RESULT 2  
US-07-683-957B-7  
; Sequence 7, Application US/07683957B  
; Patent No. 5310880



GENERAL INFORMATION:  
APPLICANT: Donahoe, Patricia K.  
APPLICANT: Ragin, Richard C.  
APPLICANT: MacLaughlin, David T.  
TITLE OF INVENTION: Purification of Milerian Inhibiting  
SUBSTANCE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/583,957B  
FILING DATE: 19910412  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldstein, Jorge A.  
REGISTRATION NUMBER: 29,021  
REFERENCE/DOCKET NUMBER: 0609.3060000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 466-0800  
TELEFAX: (202) 833-8716  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-683-957B-7

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
|||  
DB 1 RGDS 4

RESULT 3  
US-07-834-848-16  
Sequence 16, Application US/07834848  
Patent No. 5436221  
GENERAL INFORMATION:  
APPLICANT: KITAGUCHI, HIROSHI  
APPLICANT: KOMAZAWA, HIROYUKI  
APPLICANT: KOJIMA, MASAYOSHI  
APPLICANT: MORI, HIDEYO  
APPLICANT: NISHIKAWA, NAOUYUKI  
APPLICANT: SATOH, HIDEAKI  
APPLICANT: OIKASA, ATSUSHI  
APPLICANT: ONO, MITSUNORI  
APPLICANT: AZUMA, ICHIRO  
APPLICANT: SAKI, IKUO  
TITLE OF INVENTION: PEPTIDE DERIVATIVES AND APPLICATION  
THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sughrue, Lion, Zinn, Macpeak, & Seas  
STREET: 2100 Pennsy-vania Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037-3202  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/834,848  
FILING DATE: 19920213  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Biggart, Waddell A.  
REGISTRATION NUMBER: 24,861  
REFERENCE/DOCKET NUMBER: Q28480  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-834-848-16

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
|||  
DB 1 RGDS 4

RESULT 4  
US-08-127-351-54  
Sequence 54, Application US/08127351  
Patent No. 5449761  
GENERAL INFORMATION:  
APPLICANT: BELINKA Jr, BENJAMIN A.  
APPLICANT: COUGHLIN, DANIEL J.  
APPLICANT: ALVAREZ, VERNON L.  
APPLICANT: WOOD, RICHARD  
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
TITLE OF INVENTION: CONSTRUCTS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESSEE: NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,351  
FILING DATE: 28-SEP-1993  
CLASSIFICATION: 534  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-004-44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-127-351-54

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

## RESULT 5

US-08-109-106-1  
Sequence 1, Application US/08109106  
Patent No. 5475100  
GENERAL INFORMATION:  
APPLICANT: Kimikazu HASHINO et al.  
TITLE OF INVENTION: Artificial Antibody  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/109,106  
FILING DATE:

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/725,668  
FILING DATE: July 3, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:

POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:

NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:

AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:  
US-08-109-106-1

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDS 4  
Db 1 RGDS 4

## RESULT 6

US-08-251-027-8  
Sequence 8, Application US/08251027  
Patent No. 5519005

## GENERAL INFORMATION:

APPLICANT: Lider, Ofer  
APPLICANT: Greenspoon, No. 5519005m  
APPLICANT: Hershkovich, Rami  
APPLICANT: Alon, Ronen  
TITLE OF INVENTION: A METHOD OF INHIBITION OF CELLULAR AND  
MOLECULAR LEVEL BIOLOGICAL INTERACTIONS UTILIZING  
NON-PEPTIDIC SURROGATES OF THE ARG-GLY-ASP SEQUENCE

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kohn & Associates  
STREET: 30500 No. 5519005thwestern Highway, Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/251,027

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,955

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 539-5050

TELEFAX: (810) 539-5055

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

; MOLECULE TYPE: peptide  
US-08-251-027-8

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
Db 1 RGDS 4

## RESULT 7

US-08-480-367B-54  
; Sequence 54, Application US/08480367B  
; Patent No. 5578288

## GENERAL INFORMATION:

; APPLICANT: BELINKA JR, BENJAMIN A.  
; APPLICANT: COUGHLIN, DANIEL J.  
; APPLICANT: ALVAREZ, VERNON L.  
; APPLICANT: WOOD, RICHARD

; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE

; TITLE OF INVENTION: CCSTRUCTS

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LOWE, PRICE, LeBLANC & BECKER

; STREET: 99 Canal Center Plaza, Suite 300

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/480,367B

; FILING DATE: 07-06-95

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Villacorta, Gilberto M.

; REGISTRATION NUMBER: 34,038

; REFERENCE/DOCKET NUMBER: 2654-002A

; TELEPHONE: (703) 684-1111

; TELEFAX: (703) 684-1124

; TELEX:

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-480-367B-54

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
Db 1 RGDS 4

## RESULT 8

US-08-487-221A-54

; Sequence 54, Application US/08487221A

; Patent No. 5593656

; GENERAL INFORMATION:

; APPLICANT: BELINKA JR, BENJAMIN A.

; APPLICANT: COUGHLIN, DANIEL J.

; APPLICANT: ALVAREZ, VERNON L.

; APPLICANT: WOOD, RICHARD  
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
; TITLE OF INVENTION: CCSTRUCTS  
; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &

; ADDRESSEE: NEUSTADT, P.C.

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/487,221A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/127,351

; FILING DATE: 28-SRP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Villacorta, Gilberto M.

; REGISTRATION NUMBER: 34,038

; REFERENCE/DOCKET NUMBER: 4980-004-44

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-1000

; TELEFAX: (703) 413-2220

; TELEX: 248955 OPAT UR

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-487-221A-54

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
Db 1 RGDS 4

## RESULT 9

US-08-480-370-54

; Sequence 54, Application US/08480370

; Patent No. 5609847

; GENERAL INFORMATION:

; APPLICANT: BELINKA JR, BENJAMIN A.

; APPLICANT: COUGHLIN, DANIEL J.

; APPLICANT: ALVAREZ, VERNON L.

; APPLICANT: WOOD, RICHARD

; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &

; ADDRESSEE: NEUSTADT,

; ADDRESSEE: P.C.

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,370  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/127,351  
 FILING DATE: 28-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Villacorta, Gilberto M.  
 REGISTRATION NUMBER: 34,038  
 REFERENCE/DOCKET NUMBER: 4980-334-44  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEFAX: (703) 413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 54:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: Peptide  
 US-08-480-370-54

Query Match 100.0%; Score 21; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 Db 1 RGDS 4

RESULT 10  
 US-08-389-005-1  
 Sequence 1, Application: US/08389005  
 Patent No. 5618917  
 GENERAL INFORMATION:  
 APPLICANT: TOBACK, Gary F.  
 APPLICANT: LIESKE, John C.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING  
 TITLE OF INVENTION: AND TREATING KIDNEY DISEASES ASSOCIATED WITH ADHESION OF  
 TITLE OF INVENTION: CRYSTALS TO KIDNEY CELLS  
 NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BRINKS HOFER GILSON & LIONE  
 STREET: P.O. Box 10395  
 CITY: Chicago  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60610  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/389,005  
 FILING DATE: 15-FEB-1995  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Martin, Alice O.  
 REGISTRATION NUMBER: 35,601  
 REFERENCE/DOCKET NUMBER: 7814/4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 321-4200  
 TELEFAX: (312) 321-4299  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 FRAGMENT TYPE: internal  
 US-08-389-005-1  
 Query Match 100.0%; Score 21; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDS 4  
 Db 1 RGDS 4

RESULT 11  
 US-08-178-482-11  
 Sequence 11, Application US/08178482  
 Patent No. 5629294  
 GENERAL INFORMATION:  
 APPLICANT: DIZERGA, GERE S  
 APPLICANT: RODGERS, KATHLEEN E  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING  
 TITLE OF INVENTION: ADHESION FORMATION  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ROBBINS, DALGARV, BERLINER & CARSON  
 STREET: 201 NORTH FIGUEROA STREET, FIFTH FLOOR  
 CITY: LOS ANGELES  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 90012-2628  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/178,482  
 FILING DATE: 06-JAN-1994  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/789,231  
 FILING DATE: 07-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: SPITALS, JOHN P  
 REGISTRATION NUMBER: 29,215  
 REFERENCE/DOCKET NUMBER: 1920-314  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 977-1001  
 TELEFAX: (213) 977-1003  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-178-482-11

Query Match 100.0%; Score 21; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 Db 1 RGDS 4

RESULT 12  
 US-08-299-636-35  
 Sequence 35, Application US/08299636  
 Patent No. 5659041  
 GENERAL INFORMATION:  
 APPLICANT: POLLAK, Alfred  
 APPLICANT: KIRBY, Robert A.

```

; APPLICANT: DUNN-DUPAULT, Robert
; TITLE OF INVENTION: HYDRATING-TYPE RADIONUCLIDE CHELATORS
; TITLE OF INVENTION: HYDRATING AN X3S CONFIGURATION
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,636
; FILING DATE: 02-SEP-1994
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,911
; FILING DATE: 18-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/262/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-299-636-35

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RGDS 4
DB 1 RGDS 4

```

```

RESULT 13
US-08-279-155-34
; Sequence 34, Application US/08279155
; Patent No. 5662885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODEBY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKADO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 555 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,155
; FILING DATE: 22-JUL-1994

```

```

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8074-4005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-279-155-34

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RGDS 4
DB 1 RGDS 4

```

```

RESULT 14
US-08-142-449B-13
; Sequence 13, Application US/08142449B
; Patent No. 5668104
; GENERAL INFORMATION:
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Kojima, Katsuaki
; TITLE OF INVENTION: Physiologically Active Protein and
; TITLE OF INVENTION: Hematopoietic Stem Cell Growth Agent
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nels T. Lippert, White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,449B
; FILING DATE: 24-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lippert, Nels T.
; REGISTRATION NUMBER: 25,888
; REFERENCE/DOCKET NUMBER: 1145358-304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-819-8582
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-142-449B-13

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 RGDS 4  
 Db 1 RGDS 4

RESULT 15  
 US-08-464-456-33  
 ; Sequence 33, Application US/08464456  
 ; Patent No. 5681541  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dean, Richard T  
 ; TITLE OF INVENTION: Technetium-99m Labeled Polypeptides for  
 ; TITLE OF INVENTION: Imaging  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Allegretti, Ltd.  
 ; STREET: 10 South Wacker Drive Suite 3000  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/464,456  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5681541man, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 90,1104-V  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312 715 1000  
 ; TELEFAX: 312 715 1234  
 ; TELEX: 910-221-5317  
 ; INFORMATION FOR SEQ ID NO: 33:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-464-456-33

Query Match 100.0%; Score 21; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDS 4  
 Db 1 RGDS 4

Search completed: April 16, 2004, 07:31:41  
 Job time : 13.9 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 07:13:22 ; Search time 12.9 Seconds  
(without alignments)  
16.008 Million cell updates/sec

Title: US-09-991-588B-6

Perfect score: 21  
Sequence: 1 KGDS 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB pep.\*  
3: /cgn2\_6/prodata/2/iaa/5A COMB pep.\*  
4: /cgn2\_6/prodata/2/iaa/5B COMB pep.\*  
5: /cgn2\_6/prodata/2/iaa/5A COMB pep.\*  
6: /cgn2\_6/prodata/2/iaa/5B COMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	1	US-08-463-159-1
2	21	100.0	9	1	US-07-819-362-3
3	21	100.0	9	1	US-08-179-574-3
4	21	100.0	9	5	PCT-US93-00325-3
5	21	100.0	10	2	US-08-363-276B-2
6	21	100.0	10	3	US-08-755-034-2
7	21	100.0	10	5	PCT-US95-16718-2
8	21	100.0	10	5	PCT-US96-08995-2
9	21	100.0	12	1	US-08-321-668-16
10	21	100.0	12	1	US-08-837-941-16
11	21	100.0	13	3	US-08-318-794-24
12	21	100.0	13	4	US-08-470-106-24
13	21	100.0	16	1	US-07-942-245-54
14	21	100.0	16	1	US-07-942-245-62
15	21	100.0	16	1	US-07-942-245-63
16	21	100.0	16	1	US-07-942-245-64
17	21	100.0	16	1	US-07-942-245-65
18	21	100.0	16	1	US-07-942-245-66
19	21	100.0	16	1	US-07-942-245-67
20	21	100.0	16	1	US-07-942-245-68
21	21	100.0	16	1	US-07-942-245-69
22	21	100.0	16	1	US-07-942-245-71
23	21	100.0	16	1	US-07-942-245-72
24	21	100.0	16	1	US-07-942-245-73
25	21	100.0	16	1	US-07-942-245-74
26	21	100.0	16	1	US-07-942-245-75
27	21	100.0	16	1	US-07-942-245-76

28	21	100.0	16	1	US-07-942-245-77	Sequence 77, Appl
29	21	100.0	16	1	US-07-942-245-78	Sequence 78, Appl
30	21	100.0	16	1	US-07-942-245-79	Sequence 79, Appl
31	21	100.0	16	1	US-07-942-245-80	Sequence 80, Appl
32	21	100.0	16	1	US-07-942-245-81	Sequence 81, Appl
33	21	100.0	16	1	US-07-942-245-82	Sequence 82, Appl
34	21	100.0	16	1	US-07-942-245-85	Sequence 85, Appl
35	21	100.0	16	1	US-07-942-245-87	Sequence 87, Appl
36	21	100.0	16	1	US-07-942-245-90	Sequence 90, Appl
37	21	100.0	16	1	US-07-942-245-92	Sequence 92, Appl
38	21	100.0	16	1	US-07-942-245-93	Sequence 93, Appl
39	21	100.0	16	1	US-07-942-245-117	Sequence 117, App
40	21	100.0	16	1	US-07-942-245-124	Sequence 124, App
41	21	100.0	21	1	US-07-956-848A-25	Sequence 25, Appl
42	21	100.0	21	1	US-08-471-956-25	Sequence 25, Appl
43	21	100.0	26	1	US-07-942-245-263	Sequence 263, App
44	21	100.0	26	1	US-07-942-245-264	Sequence 264, App
45	21	100.0	26	1	US-07-942-245-265	Sequence 265, App

## ALIGNMENTS

RESULT 1  
US-08-463-159-1  
; Sequence 1, Application US/08463159  
; Patent No. 5552385  
; GENERAL INFORMATION:  
; APPLICANT: Christensen, ThorKild  
; APPLICANT: Baleschmidt, Per  
; APPLICANT: Sorensen, Hans  
; APPLICANT: Olsen, Ole  
; APPLICANT: Thim, Lars  
; TITLE OF INVENTION: A PHARMACEUTICAL FORMULATION  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5552385o No. 5552385disk of No. 5552385th America, Inc.  
; STREET: 405 Lexington Avenue, 64th floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,159  
; FILING DATE: 05 JUN 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agris Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 4489.000-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-463-159-1

Query Match 100.0%; Score 21; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
DB 1 KGDS 4

```

RESULT 2
US-37-819-361-3
; Sequence 3, Application US/07819361
; Patent No. 5338663
; GENERAL INFORMATION:
; APPLICANT: Potter, Huntington
; APPLICANT: Kayyali, Usamah
; TITLE OF INVENTION: Method of Interfering With Formation of
; TITLE OF INVENTION: Alpha-Antichymotrypsin-Beta-Protein Complex, Method of
; TITLE OF INVENTION: Inhibiting Beta-Protein Function and Compounds For Use
; TITLE OF INVENTION: Therein
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/819.361
; FILING DATE: 19920113
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU93-033A
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-819-361-3
Query Match 100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
Db 4 KGDS 7

RESULT 3
US-08-179-574-3
; Sequence 3, Application US/08179574
; Patent No. 5506097
; GENERAL INFORMATION:
; APPLICANT: Huntington Potter
; APPLICANT: Usamah Kayyali
; TITLE OF INVENTION: Compounds and Methods for Inhibiting
; TITLE OF INVENTION: Therein
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/179,574
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/819,361
; FILING DATE: 13-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU90-03A3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-179-574-3
Query Match 100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
Db 4 KGDS 7

RESULT 4
PCT-US93-00325-3
; Sequence 3, Application PC/TUS9300325
; GENERAL INFORMATION:
; APPLICANT: Potter, Huntington
; APPLICANT: Kayyali, Usamah
; TITLE OF INVENTION: Compounds and Methods for Inhibiting
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00325
; FILING DATE: 19930113
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,361
; FILING DATE: 13-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU90-03AA PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US93-00325-3
Query Match 100.0%; Score 21; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;

```



Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KGDS 4

Db 4 KGDS 7

## RESULT 5

US-08-363-276B-2  
; Sequence 2, Application US/08363276B  
; Patent No. 5969109  
; GENERAL INFORMATION:  
; APPLICANT: BONA ET AL.  
; TITLE OF INVENTION: CHIMERIC ANTIBODIES  
; TITLE OF INVENTION: COMPRISING ANTIGEN BINDING SITES AND B AND T CELL EPITOPES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brumbaugh, Graves, Donohue &  
; ADDRESSEE: Raymond  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10112-0228  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/363,276B  
; FILING DATE: 22-DECEMBER-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 07/486,546  
; FILING DATE: 28-FEBRUARY-1990 (ABANDONED)  
; APPLICATION NUMBER: USSN 07/687,376  
; FILING DATE: 18-APRIL-1991 (ABANDONED)  
; APPLICATION NUMBER: USSN 08/327,636  
; FILING DATE: 24-OCTOBER-1994 (ABANDONED)  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Richard S  
; REGISTRATION NUMBER: 26,154  
; REFERENCE/DOCKET NUMBER: 29889-165/29528  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-408-2558  
; TELEFAX: 212-765-2519  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Influenza Virus  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION: HA1 hemagglutinin protein  
US-08-363-276B-2

Query Match 100.0%; Score 21; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KGDS 4

Db 5 KGDS 8

## RESULT 6

US-08-755-034-2  
; Sequence 2, Application US/08755034

; Patent No. 6204250

; GENERAL INFORMATION:

; APPLICANT: BOT and BONA

; TITLE OF INVENTION: IMMUNIZATION OF INFANTS

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brumbaugh, Graves, Donohue &

; ADDRESSEE: Raymond

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10112-0228

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/755,034

; FILING DATE: 22-NOVEMBER-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Richard S

; REGISTRATION NUMBER: 26,154

; REFERENCE/DOCKET NUMBER: 29889-165/29528

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-408-2558

; TELEFAX: 212-765-2519

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Influenza Virus

; FEATURE:

; NAME/KEY:

; LOCATION:

; OTHER INFORMATION: HA1 hemagglutinin protein

US-08-755-034-2

Query Match 100.0%; Score 21; DB 3; Length 10;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KGDS 4

Db 5 KGDS 8

## RESULT 7

PCT-US95-16718-2

; Sequence 2, Application PC/TUS9516718

; GENERAL INFORMATION:

; APPLICANT: MOUNT SINAI SCHOOL OF MEDICINE OF THE

; APPLICANT: CITY UNIVERSITY OF NEW YORK

; TITLE OF INVENTION: CHIMERIC ANTIBODIES

; TITLE OF INVENTION: COMPRISING ANTIGEN BINDING SITES AND B AND T CELL EPITOPES

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brumbaugh, Graves, Donohue &

; ADDRESSEE: Raymond

; STREET: 30 Rockefeller Plaza

; CITY: New York

10 11

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
;
; OTHER INFORMATION: /note= hu p55 TNF-R mutant
; OTHER INFORMATION: construct 176-177
US-08-321-668-16

Query Match 100.0%; Score 21; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 5 KGDS 8

RESULT 10
US-08-937-941-16
; Sequence 16, Application US/08837941
; Patent No. 5766917
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: BRAXEBUSCH, Cord
; APPLICANT: VARPOLOMEY, Eugene
; APPLICANT: BATKIN, Michael
; TITLE OF INVENTION: MOLECULES INFLUENCING THE SHEDDING OF
; TITLE OF INVENTION: THE TNF RECEPTORS, THEIR PREPARATION AND THEIR USE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,941
; FILING DATE: 28-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/321,668
; FILING DATE: 12-OCT-1994
; APPLICATION NUMBER: IL 107268
; FILING DATE: 12-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: /note= hu p55 TNF-R mutant
; OTHER INFORMATION: construct 176-177
US-08-937-941-16

Query Match 100.0%; Score 21; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 5 KGDS 8

RESULT 11
US-08-318-794-24
; Sequence 24, Application US/08318794
; Patent No. 6022726
; GENERAL INFORMATION:
; APPLICANT: Palese, Peter
; TITLE OF INVENTION: GENETICALLY ENGINEERED ATTENUATED
; TITLE OF INVENTION: VIRUSES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,794
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7682-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-318-794-24

Query Match 100.0%; Score 21; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4
DB 5 KGDS 8

RESULT 12
US-08-473-106-24
; Sequence 24, Application US/08470106
; Patent No. 6316243
; GENERAL INFORMATION:
; APPLICANT: Palese, Peter
; TITLE OF INVENTION: GENETICALLY ENGINEERED ATTENUATED
; TITLE OF INVENTION: VIRUSES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

```
/
/ COUNTRY: J.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.C, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/470,106
/ FILING DATE: 06-Jun-1995
/ CLASSIFICATION: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A.
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 7682-025
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8964
/ INFORMATION FOR SEQ ID NO: 24:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 13 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-08-470-106-24

Query Match 100.0%; Score 21; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4
Db 5 KGDS 8

RESULT 13
US-07-942-245-54
/ Sequence 54, Application US/07942245
/ Patent No. 5639641
/ GENERAL INFORMATION:
/ APPLICANT: PEDERSEN, Jan T.
/ APPLICANT: SEARLE, Stephen M.J.
/ APPLICANT: REES, Anthony R.
/ APPLICANT: ROGUSKA, Michael A.
/ APPLICANT: GUILD, Braydon C.
/ TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
/ TITLE OF INVENTION: ANTIPODIES
/ NUMBER OF SEQUENCES: 522
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
/ STREET: 2100 Pennsylvania Avenue, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: United States
/ ZIP: 20037-3202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: HP 9000/700 Workstation
/ OPERATING SYSTEM: UNIX
/ SOFTWARE: In house
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/942,245
/ FILING DATE: 09-SEP-1992
/ CLASSIFICATION: 530
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 293-7060
/ TELEFAX: (202) 293-7860
/ TELEX: 6491103
/ INFORMATION FOR SEQ ID NO: 54:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 16 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-07-942-245-62

Query Match 100.3%; Score 21; DB 1; Length 16;
Best Local Similarity 100.3%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4
Db 8 KGDS 11

RESULT 14
US-07-942-245-62
/ Sequence 62, Application US/07942245
/ Patent No. 5639641
/ GENERAL INFORMATION:
/ APPLICANT: PEDERSEN, Jan T.
/ APPLICANT: SEARLE, Stephen M.J.
/ APPLICANT: REES, Anthony R.
/ APPLICANT: ROGUSKA, Michael A.
/ APPLICANT: GUILD, Braydon C.
/ TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
/ TITLE OF INVENTION: ANTIPODIES
/ NUMBER OF SEQUENCES: 522
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
/ STREET: 2100 Pennsylvania Avenue, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: United States
/ ZIP: 20037-3202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: HP 9000/700 Workstation
/ OPERATING SYSTEM: UNIX
/ SOFTWARE: In house
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/942,245
/ FILING DATE: 09-SEP-1992
/ CLASSIFICATION: 530
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 293-7060
/ TELEFAX: (202) 293-7860
/ TELEX: 6491103
/ INFORMATION FOR SEQ ID NO: 62:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 16 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-07-942-245-62

Query Match 100.3%; Score 21; DB 1; Length 16;
Best Local Similarity 100.3%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4
Db 8 KGDS 11

RESULT 15
US-07-942-245-63
/ Sequence 63, Application US/07942245
/ Patent No. 5639641
/ GENERAL INFORMATION:
/ APPLICANT: PEDERSEN, Jan T.
/ APPLICANT: SEARLE, Stephen M.J.
/ APPLICANT: REES, Anthony R.
/ APPLICANT: ROGUSKA, Michael A.
```

? APPLICANT: GUILD, Braydon C.  
 ? TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT  
 ? TITLE OF INVENTION: ANTIPODIES  
 ? NUMBER OF SEQUENCES: 522  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
 ? STREET: 2100 Pennsylvania Avenue, N.W.  
 ? CITY: Washington  
 ? STATE: D.C.  
 ? COUNTRY: United States  
 ? ZIP: 20037-3202  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: HP 9000/700 Workstation  
 ? OPERATING SYSTEM: UNIX  
 ? SOFTWARE: In house  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/07/942,245  
 ? FILING DATE: 09-SEP-1992  
 ? CLASSIFICATION: 530  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (202) 293-7060  
 ? TELEFAX: (202) 293-7860  
 ? TELEX: 6491103  
 ? INFORMATION FOR SEQ ID NO: 63:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 16 amino acids  
 ? TYPE: amino acid  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: peptide  
 ? US-07-942-245-63

Query Match 100.0%; Score 21; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4  
 Db 8 KGDS 11

Search completed: April 16, 2004, 07:31:39  
 Job time : 13.9 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:55:57 ; Search time 43.7 Seconds  
(without alignments)  
25.862 Million cell updates/sec

Title: US-09-991-588B-6

Perfect score: 21  
Sequence: 1 KGDS 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description
1	21	100.0	Aaw06492 Growth ho
2	21	100.0	Aay80473 Cell adhe
3	21	100.0	Aab91986 Fibronect
4	21	100.0	Aab86865 Transport
5	21	100.0	Aar57233 Cyclised
6	21	100.0	Aar71638 Influenza
7	21	100.0	Aar94593 Cathepsin
8	21	100.0	Abg68625 Corneodes
9	21	100.0	Aar99504 B-cell ep
10	21	100.0	Aaw12228 Influenza
11	21	100.0	Aaw60684 Influenza
12	21	100.0	Abb79917 Influenza
13	21	100.0	Aar75021 Deletion
14	21	100.0	Abg68626 Corneodes
15	21	100.0	Aaw86762 Cyclic pe
16	21	100.0	Aag78910 Human per
17	21	100.0	Aar52108 Mouse lig
18	21	100.0	Aar52101 Mouse lig
19	21	100.0	Aar52093 Mouse lig
20	21	100.0	Aar52121 Mouse lig
21	21	100.0	Aar52095 Mouse lig
22	21	100.0	Aar52104 Mouse lig
23	21	100.0	Aar52114 Mouse lig
24	21	100.0	Aar52153 Mouse lig
25	21	100.0	Aar52102 Mouse lig

26	21	100.0	16	2	AAR52083	Mouse lig
27	21	100.0	16	2	AAR52092	Mouse lig
28	21	100.0	16	2	AAR52096	Mouse lig
29	21	100.0	16	2	AAR52107	Mouse lig
30	21	100.0	16	2	AAR52109	Mouse lig
31	21	100.0	16	2	AAR52119	Mouse lig
32	21	100.0	16	2	AAR52103	Mouse lig
33	21	100.0	16	2	AAR52091	Mouse lig
34	21	100.0	16	2	AAR52094	Mouse lig
35	21	100.0	16	2	AAR52098	Mouse lig
36	21	100.0	16	2	AAR52105	Mouse lig
37	21	100.0	16	2	AAR52111	Mouse lig
38	21	100.0	16	2	AAR52100	Mouse lig
39	21	100.0	16	2	AAR52106	Mouse lig
40	21	100.0	16	2	AAR52146	Mouse lig
41	21	100.0	16	2	AAR52122	Mouse lig
42	21	100.0	16	2	AAR52116	Mouse lig
43	21	100.0	16	2	AAR52097	Mouse lig
44	21	100.0	16	2	AAR52110	Mouse lig
45	21	100.0	16	2	AAR52161	Cyclic pe

ALIGNMENTS

RESULT 1

AAW06492  
ID AAW06492 standard; peptide; 4 AA.

XX AAW06492;

DT 05-FEB-1997 (first entry)

DE Growth hormone stabilising peptide.

KW Stabilise; growth hormone; human growth hormone; oxidation; deamination.

OS Synthetic.

XX US552385-A.

XX 03-SEP-1996.

PP 05-JUN-1995; 95US-00463159.

PR 05-JUN-1995; 95US-00463159.

XX (NOVO ) NOVO-NORDISK AS.

XX Olsen O, Thim L, Balschmidt P, Christensen T, Sorensen H;

XX WPI; 1996-412094/41.

XX Stabilised growth hormone formulations - contg. tetra:peptide stabiliser.

XX Claim 1; Col 7; 5pp; English.

XX This sequence represents a peptide which acts, in a formulation, to stabilise a growth hormone, pref human growth hormone. The peptide acts to stabilise the formulation against deamination and oxidation. The formulation is stable at pH 6.8. The peptide is added to the formulation in an amount of approx. 3-5 mM

XX Sequence 4 AA;

XX Query Match 100.0%; Score 21; DB 2; Length 4;

XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGDS 4

Db 1 KGDS 4

RESULT 2  
 AAY80473  
 ID AAY80473 standard; peptide; 4 AA.  
 XX  
 AC AAY80473;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Cell adhesion peptide #8.  
 XX  
 KW Bone regenerative; osteopathic; osseous tissue; reconstitution;  
 KW scaffold matrix; bone formation promoter; bone resorption inhibitor;  
 KW cell adhesion; osteoblast; osteoclast; bone defect; fracture.  
 XX  
 OS Synthetic.  
 XX  
 FN WO200004941-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 22-JUL-1999; 99WO-USC16800.  
 XX  
 PR 24-JUL-1998; 98US-00122348.  
 XX  
 PA (PEAR-) PHARMACAL BIOTECHNOLOGIES INC.  
 XX  
 PI Budny JA;  
 XX  
 DR WPI; 2000-195084/17.  
 XX  
 XX System for reconstructing osseous tissue, useful e.g. for treating  
 PT fractures, comprises scaffold containing promoter of bone formation and  
 PT inhibitor of bone resorption.  
 XX  
 PS Claim 14; Page 31; 44pp; English.  
 XX  
 CC The invention relates to a novel system for reconstitution of osseous  
 CC tissue comprising a scaffold carrying a compound (I) that promotes bone  
 CC formation and a component that decreases bone resorption (II). (I)  
 CC induces migration and adhesion of osteoblasts and osteoclasts and (II)  
 CC inhibits proteolysis (specifically by plasmin) of extracellular matrix.  
 CC (I) is preferably selected from: selectin or selectin binding fragments,  
 CC proteins and peptides that facilitate cell adhesion, plasminogen  
 CC activator inhibitors, protease inhibitors and metalloprotease inhibitors.  
 CC The peptides AAY80466-Y80492 are claimed examples of cell adhesion  
 CC peptides used in the system of the invention. The system is used to  
 CC replace, remodel or correct bone defects, e.g. fractures, fissures or  
 CC bone mass loss. Incorporation of (I) into the scaffold results in rapid  
 CC seeding by osteoblasts and the development of an organic matrix, i.e. the  
 CC preformed scaffold replaces the rate-determining step of extracellular  
 CC matrix formation. The scaffold can be designed to have a predetermined  
 CC resorption/degradation rate, and may include regulatory compounds for  
 CC specific cell types  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 3; Length 4;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 Db 1 KGDS 4  
 RESULT 3  
 AAB91986  
 ID AAB91986 standard; peptide; 4 AA.  
 XX  
 AC AAB91986;  
 XX  
 DT 22-JUN-2000 (first entry)

XX Fibrinectin fragment and fibrin related peptide SEQ ID NO:1162.  
 DE  
 XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidy1; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200069900-A2.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 17-MAY-2000; 2000WO-US013576.  
 XX  
 PR 17-MAY-1999; 99US-0134406P.  
 PR 10-SEP-1999; 99US-0153406P.  
 PR 15-OCT-1999; 99US-0159783P.  
 XX  
 PA (CONJ-) CONJUCHEM INC.  
 XX  
 PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX  
 DR WPI; 2001-112059/12.  
 XX  
 XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity.  
 PT  
 PS Disclosure; Page 575; 733pp; English.  
 XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidy1 and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity in  
 CC vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 4; Length 4;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 Db 1 KGDS 4  
 RESULT 4  
 AAB86865  
 ID AAB86865 standard; peptide; 4 AA.  
 XX  
 AC AAB86865;  
 XX  
 DT 28-NOV-2001 (first entry)  
 XX  
 XX Transport molecule/ligand binding-associated peptide #11.  
 DE  
 XX Transport molecule; ligand; cancer treatment; autoimmune disease;  
 KW inflammation; infection.  
 KW

OS Synthetic.  
 PN WO200168142-A1.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 13-MAR-2001; 2001WO-EP0002833.  
 XX  
 PR 13-MAR-2000; 2000DE-01012120.  
 XX  
 PA (KTB)- KTB TUMFORSCHUNGS GMBH.  
 XX  
 PI Kratz F;  
 XX  
 DR WPI; 2001-589998/66.  
 XX  
 XX New ligand, comprising therapeutic or diagnostic agent bonded non-covalently with substance having high affinity to transport molecule.  
 PT  
 PT Disclosure; Page 39; 74pp; German.  
 XX  
 PS This invention describes novel ligands which bind to transport molecules, comprising a therapeutic and/or diagnostic agent (A) non-covalently bonded via a linkage cleavable in vivo depending on pH and/or enzymatically with a substance (B) having an association constant KA to a transport molecule of above 10<sup>3</sup> M<sup>-1</sup>, is new. The medicaments are especially useful for the treatment of cancers, autoimmune diseases, acute and chronic inflammation and infections caused by viruses or microorganisms. The diagnostic kits are useful for the detection of these illnesses and for the detection of the transport molecule and/or its distribution in vivo. The ligands have excellent solubility in the medium at the site of action and are easy and inexpensive to convert into adducts, as the interaction with the transport material is physical.  
 CC AAB6843-AAB8620 represent peptides used to illustrate the method of the invention  
 CC  
 XX  
 SQ Sequence 4 AA;  
 Query Match 100.0%; Score 21; DB 4; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KGDS 4  
 Db 1 KGDS 4  
 RESULT 5  
 AAR57233  
 ID AAR57233 standard; peptide; 7 AA.  
 XX AC AAR57233;  
 XX 25-MAR-2003 (revised)  
 DT 17-APR-1995 (first entry)  
 XX  
 DE Cyclised integrin receptor antagonist.  
 XX Integrin receptor antagonist; cyclic;  
 KW cell adhesion modulation. fibronectin.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT Modified-site 2 /note= "nicotinyl-Lys"  
 FT  
 XX WO9415958-A2.  
 PN  
 XX 21-JUL-1994.  
 PD  
 XX 07-JAN-1994; 94WC-IB000026.  
 PF

XX 08-JAN-1993; 93US-00001773.  
 PR  
 XX (TANA) TANABE SEIYAKU CO.  
 XX  
 PI Chiang SN, Cardarelli PM, Lobl TJ;  
 XX  
 DR WPI; 1994-249134/30.  
 XX  
 XX Cyclised integrin receptor antagonists - useful for modulating cell adhesion and no therapeutic agents.  
 PT  
 PT Table 3; Page 66; 99pp; English.  
 PS  
 XX The invention relates to new cyclised peptides which act as integrin receptor antagonists. The peptides are described in a highly generic formula, and they do not contain the RGD motif. The peptides modulate cell adhesion activity. They have IC50 values of < 500 (sometimes < 10) micromolar in the U937-fibronectin adhesion assay. They are useful therapeutically for treating cell adhesion-associated inflammatory diseases such as rheumatoid arthritis, asthma, allergy, ARDS and CC inflammatory bowel syndrome; autoimmune disease; cardiovascular disease CC (e.g. thrombosis); neoplastic disease; and as contraceptives, or for CC promoting organ transplants or wound healing. The present sequence is a CC specific example of the generic peptides. (Updated on 25-MAR-2003 to CC correct PN field.)  
 XX  
 SQ Sequence 7 AA;  
 Query Match 100.0%; Score 21; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KGDS 4  
 Db 2 KGDS 5  
 RESULT 6  
 AAR71638  
 ID AAR71638 standard; peptide; 9 AA.  
 XX AC AAR71638;  
 XX 25-MAR-2003 (revised)  
 DT 20-SEP-1995 (first entry)  
 XX  
 DE Influenza hemagglutinin antigenic site B.  
 XX HIV-1; gp41; gp160; 2F5 epitope; antibody; cell fusion; AIDS; vaccine;  
 KW immunogen; therapy; human immunodeficiency virus type 1;  
 KW monoclonal antibody; Mab; hemagglutinin.  
 XX Influenza virus.  
 OS  
 XX WO9507354-A1.  
 PN  
 XX 16-MAR-1995.  
 PD  
 XX 12-SEP-1994; 94WO-EP003039.  
 PF  
 XX 11-SEP-1993; 93EP-00114631.  
 PR  
 XX (POLY-) POLYMUN SCI IMMUNOBIOLOGISCHE FORSCHUNG.  
 PA  
 XX Katinger H, Muster T;  
 PI  
 XX WPI; 1995-123428/16.  
 DR  
 DR N-PSDB; AAQ86043.  
 XX  
 XX New peptide(s) corresponding to a HIV-1 gp 160 epitope - used to elicit PT antibodies which neutralise different HIV-1 strains and inhibit cell PT fusion caused by HIV-1.



XX PS Disclosure; Fig 1; 28pp; English.

XX CC The peptides given in AAR71631-32 are based on amino acids 661-668 of HIV

CC -1 isolate RH10 gp41, the epitope of human MAB 2P5. The peptides were

CC inserted into antigenic site B of influenza virus WSN hemagglutinin

CC (AAR71638), by gene fusion, to give the constructs HA-ELDKWAS (AAR71639)

CC and HA-LELDKXAS (AAR71640). Sera from mice immunized with the chimeric

CC influenza viruses showed neutralization activity against HIV-1 MN, RF and

CC IIB. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4

Db 3 KGDS 6

RESULT 7

ID AAR94593 standard; peptide; 9 AA.

XX AC AAR94593;

XX DT 25-MAR-2003 (revised)

DT 21-AUG-1996 (first entry)

XX DE Cathepsin G protease active site sequence.

XX KW Beta-amyloid; Alzheimer's disease; cholinesterase; lipase; ebelactone-A;

KW serine protease; para-aminophenylmethanesulphonyl fluoride; inhibition;

KW complex formation; alpha(1)-antichymotrypsin; Down's diseases; ageing;

KW cytotoxic T cell protease; cathepsin.

XX OS Synthetic.

XX US US5506097-A.

XX 09-APR-1996.

XX 10-JAN-1994; 94US-00179574.

XX 24-AUG-1990; 90US-00572571.

PR 13-JAN-1992; 92US-00819361.

PR 13-JAN-1993; 93WO-US000325.

XX PA (HARD ) HARVARD COLLEGE.

XX Kayyali U, Potter H;

XX WPI; 1996-200270/20.

XX Inhibiting enzymatic activity of Alzheimer amyloid beta-protein - using p

PT -aminophenylmethanesulphonyl fluoride or ebelactone A, for treatment,

PT study and diagnosis of Alzheimer's disease, etc.

XX Disclosure; Fig 1; 17pp; English.

XX A novel method of inhibiting the esterase (cholinesterase and lipase)

CC activities of the beta-amyloid protein associated with Alzheimer's

CC disease. The beta-amyloid protein contains an esterase activity deduced

CC based on active site similarities with serine proteases (see AAR94592-

CC 96). The esterase activity of the beta-amyloid protein is inhibited by

CC the cpds. of the invention i.e. ebelactone A or para-

CC amidophenylmethanesulphonyl fluoride. Inhibition of these activities

CC prevent complex formation between the beta-amyloid protein and alpha(1)-

CC antichymotrypsin, thus can be used to treat, study or diagnose

CC Alzheimer's or Down's diseases or normal ageing. (Updated on 25-MAR-2003

CC to correct PF field.)

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4

Db 4 KGDS 7

RESULT 8

ID ABG68625 standard; peptide; 9 AA.

XX AC ABG68625;

XX DT 07-OCT-2002 (first entry)

XX DE Corneodesmosin peptide #6.

XX KW Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;

KW stratum corneum; chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCF;

KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;

KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;

KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;

KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;

KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;

KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;

KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;

KW antipsoriatic; dermatological; antiinflammatory; anti-allergic.

XX OS Unidentified.

XX WO WO200244736-A2.

XX 06-JUN-2002.

XX 30-NOV-2001; 2001WO-GB005303.

XX 30-NOV-2000; 2000GB-00029225.

PR 07-DEC-2000; 2000GB-00029879.

XX (MOLE-) MOLECULAR SKINCARE LTD.

XX Tazi-Ahmini R, Bavić C, Ward S, Duff G, Cork M;

XX WPI; 2002-557554/59.

XX Diagnosing a disease or susceptibility to a disease associated with

PT abnormal cell-cell adhesion between epithelial cells, by detecting

PT mutation in nucleic acid encoding adhesion protein, protease or protease

PT inhibitor.

XX Example A5; Page 102; 257pp; English.

XX The invention relates to a method for diagnosis of a disease or

CC susceptibility to a disease associated with abnormal cell-cell adhesion

CC between epithelial cells, comprising detecting a mutation in a nucleic

CC acid encoding an adhesion protein, a protease or a protease inhibitor, or

CC modulated level of adhesion protein, protease or protease inhibitor, or

CC its fragment polypeptide. The method is useful for diagnosing a disease

CC such as eczema, contact dermatitis, lung atopic asthma, post viral

CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary

CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic

CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous

CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,

CC malignancy of the gastrointestinal tract, malignancy of the lung,

CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents

CC a peptide used in the scope of the invention

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4  
 ||||  
 Db 3 KGDS 6

## RESULT 9

AAR99504  
 ID AAR99504 standard; peptide; 10 AA.  
 XX  
 AC AAR99504;  
 XX  
 DT 09-JAN-1997 (first entry)  
 XX  
 DE B-cell epitope used in construction of chimeric immunoglobulin.  
 XX  
 KW Chimera; chimeric; immunoglobulin; Ig; B-cell; T-cell; lymphocyte;  
 KW epitope; immune response; vaccine; pathogen; antibody; influenza;  
 KW measles; hepatitis; foot and mouth disease; tetanus toxoid;  
 KW human immunodeficiency virus; HIV; heat shock protein; M protein;  
 KW hen egg white lysozyme; nuclease.  
 XX  
 OS Influenza virus.

XX WO9619584-A1.

PN 27-JUN-1996.

PP 21-DEC-1995; 95WO-US016713.

PR 22-DEC-1994; 94US-00363276.

XX (MOUN ) MOUNT SINAI SCHOOL MEDICINE.

XX Bona C, Zaghouani H;

XX WPI; 1996-309598/31.

XX Chimeric immunoglobulin with CDR loop subst.d. for T and or B cell epitope  
 PT - useful in vaccine composition to enhance immune response to pathogens.

XX Disclosure; Page 17; 131pp; English.

XX Chimeric immunoglobulins (Ig) having a CDR loop of the parent Ig replaced  
 CC with a foreign peptide sequence corresponding to a T- or B- cell epitope,  
 CC may be used in vaccine compositions to enhance an immune response to a  
 CC pathogen. Chimeric Ig comprising a B-cell epitope can also be used to  
 CC label B-cells, to test the ability of a subject to mount a humoral  
 CC response to a particular B-cell epitope or to collect B-cells which  
 CC recognise the epitope. . An antibody comprising a chimeric Ig molecule  
 CC which comprises an antigen binding site may be used in diagnostic assays  
 CC to detect the presence of a particular target antigen, which binds to the  
 CC antibody binding site. Sequences of B-cell epitopes are given in AAR99503  
 CC -08. Sequences of T-cell epitopes are given in AAR99509-17. The term B-  
 CC cell epitope refers to a peptide which is able to bind to an  
 CC immunoglobulin receptor of a B-cell and participate in the induction of  
 CC antibody production by that B-cell. This peptide is an immunodominant B-  
 CC cell epitope in site B of influenza virus HA1 haemagglutinin

XX Sequence 10 AA;

Query Match 100.0%; Score 21; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4  
 ||||  
 Db 5 KGDS 8

## RESULT 10

AAM12228  
 ID AAM12228 standard; peptide; 10 AA.

XX  
 AC AAM12228;

DT 06-AUG-1997 (first entry)

XX Influenza HA1 haemagglutinin B-cell epitope.

XX B-cell epitope; gp120; hypervariable region 3 loop; envelope protein;  
 KW immunoglobulin; vaccine; cancer; viral infection; leukaemia; V3 loop;  
 KW influenza virus; target antigen; helper T-cell epitope; immune response;  
 KW cell-surface receptor; HIV-1; immune system; therapy.

XX Haemophilus influenzae.

XX WO9636357-A1.

XX 21-NOV-1996.

XX 13-MAY-1996; 96WO-US006756.

XX 15-MAY-1995; 95US-00441328.

XX 07-JUN-1995; 95US-00477424.

XX (BONA/) BONA C A.

XX (LEEY/) LEE Y C.

XX (BRUM/) BRUMEANU T.

XX (DEHA/) DEHAZYA P.

XX Bona CA, Lee YC, Brumeanu T, Dehazy P;

XX WPI; 1997-011851/01.

XX New methods of coupling peptide(s) to immuno-globulin(s) via carbohydrate

XX - and related conjugates, useful for treating cancer and viral infections

XX and in vaccines.

XX Disclosure; Page 10; 69pp; English.

XX AAM1227-W12232 represent B-cell epitopes that can be used in the method  
 CC of the invention. The method of the invention is for conjugating a  
 CC peptide (such as this sequence) to an immunoglobulin (Ig) molecule, via a  
 CC carbohydrate (CB) residue of the Ig. The method comprises enzymatically  
 CC oxidising the CB residue of the Ig, then reacting this residue with an  
 CC amino group of the peptide. The product of this reaction is then  
 CC stabilised by reacting it with a reducing agent. The conjugate may be  
 CC used as a vaccine and in diagnostic and therapeutic methods. The  
 CC conjugates are used to treat cancer and viral infections, such as  
 CC leukaemia and influenza viruses. The conjugates are used to obtain a  
 CC ligand of the peptide using affinity purification, and to detect and  
 CC quantitate the amount of a target antigen of the Ig. When the peptide is  
 CC a B-cell epitope, the conjugates are used to label B-cells, and when the  
 CC peptide is a helper T-cell epitope, the conjugates are used to test the  
 CC ability of a subject to mount an immune response to that epitope. The  
 CC conjugates are more effective in eliciting an immune response than  
 CC unconjugated peptides. The conjugates increase the half-life of the  
 CC peptide and also, via binding of the FC region of the Ig to cell-surface  
 CC receptor, recruit elements of the immune system to improve the overall  
 CC efficiency of the immune response

XX Sequence 10 AA;

Query Match 100.0%; Score 21; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGDS 4  
 ||||  
 Db 5 KGDS 8

RESULT 11  
AAM60684  
D D AAM60684 standard; peptide; 10 AA.  
XX AC  
XAC AAM60684;  
DN 22-SEP-1998 (first entry)  
DT Influenza virus HA1 hemagglutinin protein B cell peptide epitope.  
XX Immunisation; target antigen; epitope; inoculation; infant mammal;  
XX viral antigen; depressed humoral response; respiratory syncytial virus;  
KW rotavirus; measles virus; human immunodeficiency virus; hepatitis virus;  
KW herpes simplex virus; influenza virus; Streptococcus pneumoniae;  
KW Hemophilus influenzae; Neisseria meningitidis; Staphylococcus aureus;  
KW protozoan antigen; malaria.  
XX  
OS Influenza virus.  
XX WO9322145-A1.  
XX 28-MAY-1998.  
XX 21-NOV-1997; 97WO-US021687.  
XX 22-NOV-1996; 96US-00755034.  
PP (MOUN ) MOUNT SINAI SCHOOL MEDICINE.  
PI Bot A, Bona C;  
DR WPI; 1998-312182/27.  
XX Immunisation of infant mammals - by inoculating the mammal with a nucleic acid encoding a relevant epitope of a target antigen.  
PT Disclosure; Page 3; 83pp; English.  
XX Sequence shown in AAM60683 to AAM60700 are epitope sequences of various viral antigens used to exemplify the method of invention of immunising an infant mammal against a target antigen. The method comprises inoculating the mammal with a nucleic acid encoding a relevant epitope of a target antigen in a carrier, such that the relevant epitope is expressed in the infant mammal. The genetic immunisation of infant mammals can give rise to effective cellular (including the induction of cytotoxic T lymphocytes) and humoral immune responses against the target antigen. The methods are particularly used for treating infants with depressed humoral responses, that have high-zone tolerances against the target antigens or have a Th2 biased immune response. The target antigen may be a viral antigen, e.g. a respiratory syncytial virus antigen, a rotavirus antigen, a measles virus antigen, a human immunodeficiency virus antigen, a hepatitis virus antigen, a hepatitis B virus antigen, a herpes simplex virus antigen or an influenza virus antigen, a bacterial antigen e.g. Streptococcus pneumoniae antigen, Hemophilus influenzae antigen, Neisseria meningitidis antigen, Staphylococcus aureus antigen or a protozoan antigen such as a malaria antigen.

Query March 100.0%; Score 21; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.7e+32;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGDS 4  
|||  
DB 5 KGDS 8

RESULT 12  
ABB79917  
ID ABB79917 standard; peptide; 10 AA.

KW phorbol myristate acetate; PMA; spacer region.  
 XX Synthetic.  
 OS  
 XX AU9475742-A.  
 PN  
 XX 04-MAY-1995.  
 PD  
 XX  
 XX  
 PF 11-OCT-1994; 94AU-00075742.  
 XX  
 XX 12-OCT-1993; 93IL-00107268.  
 PR  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 PA  
 XX Wallach D, Brakebusch C, Varfolomeev E, Batkin M;  
 XX WPI; 1995-134342/26.  
 XX  
 XX New protease capable of cleaving soluble tumour necrosis factor (TNF)  
 PT receptor - from cell-bound TNF- receptor, useful for antagonising  
 PT deleterious effects of TNF.  
 XX  
 XX Example 2; Fig 5; 40pp; English.  
 PS  
 XX The sequences represented in AAR75013-25 are fragments of deletion  
 XX mutants of the spacer region of human p55 tumour necrosis factor (TNF-R),  
 CC shown in AAR75012. Expression of this receptor is regulated by shedding  
 CC of the extracellular receptor fragment. The p55 TNF-R can be shed in  
 CC response to different inducing agents, e.g. phorbol myristate acetate  
 CC (PMA), depending on cell type. The only region of the receptor whose  
 CC structure affects the shedding response is the spacer region in the  
 CC extracellular domain. This region is located close to a site of cleavage  
 CC of the molecule, and links the Cys rich module to the transmembrane  
 CC domain. The sequences shown in AAR75026-47 are fragments of replacement  
 CC mutations of the spacer region. These mutations were introduced in order  
 CC to create an inhibitor of a protease that is capable of cleaving the  
 CC soluble TNF-R from the cell bound TNF-R. The relevant fragments of the  
 CC successful inhibitors can be seen in AAR75017-9, AAR75025, AAR75033-5 and  
 CC AAR75042-3. The protease inhibitors can be used for enhancing TNF  
 CC function  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. NO. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 Db |||||  
 5 KGDS 8  
 RESULT 14  
 ABG68626  
 ID ABG68626 standard; peptide; 14 AA.  
 XX  
 AC ABG68626;  
 XX  
 XX 07-OCT-2002 (first entry)  
 DT  
 XX Corneodesmosin peptide #7.  
 DE  
 XX Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;  
 KW stratum corneum chymotryptic enzyme; envoplatin; desmoplakin; pig; SCCE;  
 KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;  
 KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;  
 KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;  
 KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;  
 KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;  
 KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;  
 KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;  
 KW anti-psoriatic; dermatological; antiinflammatory; antiallergic.  
 XX

OS Unidentified.  
 XX WO200244736-A2.  
 PN  
 XX 06-JUN-2002.  
 PD  
 XX  
 XX 30-NOV-2001; 2001WO-GB005303.  
 PF  
 XX 30-NOV-2000; 2000GB-00029225.  
 PR  
 XX 07-DEC-2000; 2000GB-00029879.  
 XX  
 XX (MOLE-) MOLECULAR SKINCARE LTD.  
 PA  
 XX Tasi-Ahmini R, Bavik C, Ward S, Duff G, Cork M;  
 XX WPI; 2002-557554/59.  
 XX  
 XX Diagnosing a disease or susceptibility to a disease associated with  
 PT abnormal cell-cell adhesion between epithelial cells, by detecting  
 PT mutation in nucleic acid encoding adhesion protein, protease or protease  
 PT inhibitor.  
 XX  
 XX Example A6; Page 185; 257pp; English.  
 PS  
 XX The invention relates to a method for diagnosis of a disease or  
 CC susceptibility to a disease associated with abnormal cell-cell adhesion  
 CC between epithelial cells, comprising detecting a mutation in a nucleic  
 CC acid encoding an adhesion protein, a protease or a protease inhibitor, or  
 CC modulated level of adhesion protein, protease or protease inhibitor, or  
 CC its fragment polypeptide. The method is useful for diagnosing a disease  
 CC such as eczema, contact dermatitis, lung atopic asthma, post viral  
 CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary  
 CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic  
 CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous  
 CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,  
 CC malignancy of the gastrointestinal tract, malignancy of the lung,  
 CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents  
 CC a peptide used in the scope of the invention  
 XX  
 XX Sequence 14 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KGDS 4  
 Db |||||  
 10 KGDS 13  
 RESULT 15  
 AAW86762  
 ID AAW86762 standard; peptide; 15 AA.  
 XX  
 AC AAW86762;  
 XX  
 XX 26-MAR-1999 (first entry)  
 DT  
 XX Cyclic peptide which restores p53 protein activity to p53 mutants.  
 DE  
 XX p53; mutant; cyclic; breast cancer; lung; colon; antitumour.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 XX Modified-site 7..13  
 FT /note= "the side chain carboxy of Asp(7) is condensed  
 FT with the side-chain amino group of Lys(13) to form a  
 FT lactam bridge, thereby determining the cyclic  
 FT conformation of the peptide"  
 FT Modified-site 15  
 FT /note= "Leu-NH2"  
 XX

PN WO9851707-A1.  
 XX  
 PD 19-NOV-1998.  
 XX  
 PF 15-MAY-1999; 98WO-JP002148.  
 XX  
 PR 15-MAY-1997; 97JP-00126113.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Shibata K, Yamasaki M, Yoshida T, Mizukami T, Shinkai A;  
 PI Anazawa H;  
 XX  
 DR WPI; 1999-070115/56.  
 XX  
 PT Peptides which restore the p53 protein activity - useful for treatment of  
 PT cancer, e.g. breast, colon and lung.  
 XX  
 PS Claim 15; Page 15; 101pp; Japanese.  
 XX  
 CC New peptides are disclosed which have the effect of restoring p53 protein  
 CC activity and DNA binding activity to p53 mutant forms, such as those  
 CC found in tumour cells. The peptides may be used for the treatment of  
 CC cancers including breast, colon and lung cancer. The present sequence  
 CC represents a specific example of the new peptides  
 XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 21; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred.No. 4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KGDS 4  
 Db 5 KGDS 8  
 Search completed: April 16, 2004, 07:20:52  
 Job time : 45.7 secs